

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
6 December 2001 (06.12.2001)

PCT

(10) International Publication Number
WO 01/92891 A2

(51) International Patent Classification⁷: G01N 33/68

Robert, R.; 3309 South 116th Street, Omaha, NE 68144 (US). JOHNSON, Mark, L.; 16216 N. Circle, Omaha, NE 68135 (US).

(21) International Application Number: PCT/US01/16946

(22) International Filing Date: 25 May 2001 (25.05.2001)

(74) Agents: REA, Teresa, Stanek; Burns, Doane, Swecker & Mathis, LLP, P.O. Box 1404, Alexandria, VA 22313-1404 et al. (US).

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/578,900 26 May 2000 (26.05.2000) US

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(71) Applicants: GENOME THERAPEUTICS CORPORATION [US/US]; 100 Beaver Street, Waltham, MA 02453 (US). CREIGHTON UNIVERSITY SCHOOL OF MEDICINE [US/US]; Suite 1601, 1601 North 30th Street, Omaha, NE 68178 (US).

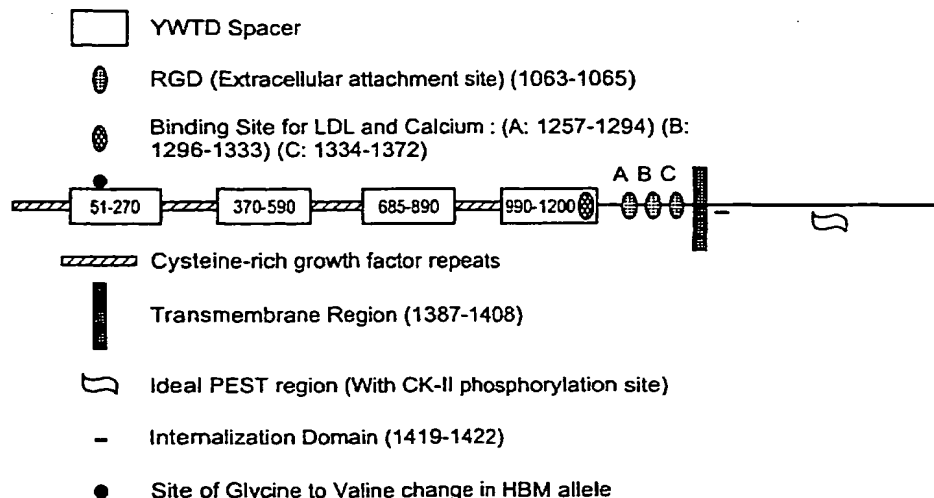
(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(72) Inventors: CARULLI, John, P.; 9 Harris Drive, Southboro, MA 01772 (US). LITTLE, Randall, D.; 73 Elm Road #6, Newtonville, MA 02460 (US). RECKER,

[Continued on next page]

(54) Title: REGULATING LIPID LEVELS VIA THE *ZMAX1* OR *HBM* GENE

Model for a LDL Receptor-Related protein, Zmax1



(57) Abstract: The present invention relates to the high bone mass (*HBM*) gene, the corresponding wild-type gene (*Zmax1*), and mutants thereof. The genes identified in the present invention are implicated in regulation of physiological lipid levels, and thereby lipid-mediated diseases and conditions. The invention also provides nucleic acids, including coding sequences, oligonucleotide primers and probes, proteins, cloning vectors, expression vectors, transformed hosts, methods of developing pharmaceutical compositions, methods of identifying molecules involved in lipid level regulation in a subject. In preferred embodiments, the present invention is directed to methods for treating and preventing atherosclerosis, arteriosclerosis cardiovascular disease, atherosclerotic and arteriosclerotic associated conditions.



Published:

— without international search report and to be republished
upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

REGULATING LIPID LEVELS VIA THE *ZMAX1* OR *HBM* GENE

INVENTORS: John P. Carulli, Randall D. Little, Robert R. Recker and Mark L. Johnson

RELATED APPLICATIONS

This application is a continuation-in-part of Application No. 09/543,771 filed April 5, 2000 and Application No. 09/544,398 filed April 5, 2000, which are continuation-in-part
5 applications of Application No. 09/229,319, filed January 13, 1999, which claims benefit of U.S. Provisional Application No. 60/071,449, filed January 13, 1998, and U.S. Provisional Application No. 60/105,511, filed October 23, 1998, all of which are herein incorporated by reference in their entirety.

FIELD OF THE INVENTION

10 The present invention relates generally to the field of genetics, genomics and molecular biology. More particularly, the invention relates to methods and materials used to isolate, detect and sequence a high bone mass gene and corresponding wild-type gene, and
mutants thereof that may be involved with modulating lipid levels. The present invention also relates to the high bone mass gene, the corresponding wild-type gene, and mutants
15 thereof. The genes identified in the present invention are implicated in the ontology and physiology of atherosclerosis, arteriosclerosis and associated diseases and conditions related thereto. The invention also provides nucleic acids, proteins, cloning vectors, expression
vectors, transformed hosts, methods of developing pharmaceutical compositions, methods of identifying molecules involved in arteriosclerosis and associated conditions, and methods of
20 treating or preventing diseases associated with abnormal lipid levels. In preferred embodiments, the present invention is directed to methods for treating, diagnosing,

preventing and screening for normal and abnormal lipid-associated conditions, including arteriosclerosis, cardiovascular disease and stroke.

BACKGROUND OF THE INVENTION

5 Cardiovascular disease is the number one killer in the United States, and atherosclerosis is the major cause of heart disease and stroke. It is widely appreciated that cholesterol plays an important role in atherogenesis. Normally, most cholesterol serves as a structural element in the walls of cells, whereas much of the rest is in transit through the blood or functions as the starting material for the synthesis of bile acids in the liver, steroid
10 hormones in endocrine cells and vitamin D in skin. The transport of cholesterol and other lipids through the circulatory system is facilitated by their packaging into lipoprotein carriers. These spherical particles comprise protein and phospholipid shells surrounding a core of neutral lipid, including unesterified ("free") or esterified cholesterol and triglycerides. Risk for atherosclerosis increases with increasing concentrations of low density lipoprotein (LDL)
15 cholesterol, whereas risk is inversely proportional to the levels of high density lipoprotein (HDL) cholesterol. The receptor-mediated control of plasma LDL levels has been well-defined, and recent studies have now provided new insights into HDL metabolism.

 The elucidation of LDL metabolism began in 1974 by Michael Brown and Joseph Goldstein. In brief, the liver synthesizes a precursor lipoprotein (very low density
20 lipoprotein, VLDL) that is converted during circulation to intermediate density lipoprotein (IDL) and then to LDL. The majority of the LDL receptors expressed in the body are on the surfaces of liver cells, although virtually all other tissues ("peripheral tissues") express some LDL receptors. After binding, the receptor-lipoprotein complex is internalized by the cells via coated pits and vesicles, and the entire LDL particle is delivered to lysosomes, wherein it

is disassembled by enzymatic hydrolysis, releasing cholesterol for subsequent cellular metabolism. This whole-particle uptake pathway is called "receptor-mediated endocytosis." Cholesterol-mediated feedback regulation of both the levels of LDL receptors and cellular cholesterol biosynthesis help ensure cellular cholesterol homeostasis. Genetic defects in the

5 LDL receptor in humans results in familial hypercholesterolemia, a disease characterized by elevated plasma LDL cholesterol and premature atherosclerosis and heart attacks. One hypothesis for the deleterious effects of excess plasma LDL cholesterol is that LDL enters the artery wall, is chemically modified, and then is recognized by a special class of receptors called macrophage scavenger receptors, that mediate the cellular accumulation of the LDL

10 cholesterol in the artery, eventually leading to the formation of an atherosclerotic lesion.

The major lipoprotein classes include intestinally derived chylomicrons that transport dietary fats and cholesterol, hepatic-derived VLDL, IDL and LDL that can be atherogenic, and hepatic- and intestinally-derived HDL that are antiatherogenic. Apoprotein B (ApoB) is necessary for the secretion of chylomicrons (Apo B48) and VLDL, IDL and LDL (Apo

15 B100). Plasma levels of VLDL triglycerides are determined mainly by rates of secretion in LPL lipolytic activity. Plasma levels of LDL cholesterol are determined mainly by the secretion of Apo B100 into plasma, the efficacy with which VLDL are converted to LDL and by LDL receptor-mediated clearance. Regulation of HDL cholesterol levels is complex and is affected by rates of synthesis of its Apo proteins, rates of esterification of free cholesterol to

20 cholesterol ester by LCAT, levels of triglyceride-rich lipoproteins and CETP-mediated transfer of cholesterol esters from HDL, and clearance from plasma of HDL lipids and Apo proteins.

Normal lipoprotein transport is associated with low levels of triglycerides and LDL cholesterol and high levels of HDL cholesterol. When lipoprotein transport is abnormal,

lipoprotein levels can change in ways that predispose individuals to atherosclerosis and arteriosclerosis (see Ginsberg, *Endocrinol. Metab. Clin. North Am.*, 27: 503-19 (1998)).

Several lipoprotein receptors may be involved in cellular lipid uptake. These receptors include: scavenger receptors; LDL receptor-related protein/ α 2-macroglobulin receptor (LRP); LDL receptor; and VLDL receptor. With the exception of the LDL receptor, all of these receptors are expressed in atherosclerotic lesions while scavenger receptors are mostly expressed in macrophages, the LRP and VLDL receptors may play an important role in mediating lipid uptake in smooth muscle cells (Hiltunen *et al.*, *Atherosclerosis*, 137 suppl.: S81-8 (1998)).

10 A major breakthrough in the pharmacologic treatment of hypercholesterolemia has been the development of the "statin" class of 3-hydroxy-3-methylglutaryl-CoA reductase (HMG CoA reductase) inhibitory drugs. 3-Hydroxy-3-methylglutaryl-CoA reductase is the rate controlling enzyme in cholesterol biosynthesis, and its inhibition in the liver stimulates LDL receptor expression. As a consequence, both plasma LDL cholesterol levels and the risk
15 for atherosclerosis decrease. The discovery and analysis of the LDL receptor system has had a profound impact on cell biology, physiology, and medicine.

HDL is thought to remove unesterified, or "free" cholesterol (FC) from peripheral tissues, after which most of the cholesterol is converted to cholesteryl ester (CE) by enzymes in the plasma. Subsequently, HDL cholesterol is efficiently delivered directly to the liver and
20 steroidogenic tissues via a selective uptake pathway and the HDL receptor, SR-BI (class B type I scavenger receptor) or, in some species, transferred to other lipoproteins for additional transport in metabolism. For additional discussion on HDL and LDL metabolism see Krieger, *Proc. Natl. Acad. Sci. USA*, 95:4077-4080, 1998.

Recently, a strong interest in the genetic control of peak bone mass has developed in the field of osteoporosis. The interest has focused mainly on candidate genes with suitable polymorphisms to test for association with variation in bone mass within the normal range, or has focused on examination of genes and gene loci associated with low bone mass in the range found in patients with osteoporosis. The vitamin D receptor locus (VDR) (Morrison *et al.*, *Nature*, 367:284-287 (1994)), PTH gene (Howard *et al.*, *J. Clin. Endocrinol. Metab.*, 80:2800-2805 (1995); Johnson *et al.*, *J. Bone Miner. Res.*, 8:11-17 (1995); Gong *et al.*, *J. Bone Miner. Res.*, 10:S462 (1995)) and the estrogen receptor gene (Hosoi *et al.*, *J. Bone Miner. Res.*, 10:S170 (1995); Morrison *et al.*, *Nature*, 367:284-287 (1994)) have figured most prominently in this work. These studies are difficult because bone mass (the phenotype) is a continuous, quantitative, polygenic trait, and is confounded by environmental factors such as nutrition, co-morbid disease, age, physical activity, and other factors. Also, this type of study design requires large numbers of subjects. In particular, the results of VDR studies to date have been confusing and contradictory (Garnero *et al.*, *J. Bone Miner. Res.*, 10:1283-1288 (1995); Eisman *et al.*, *J. Bone Miner. Res.*, 10:1289-1293 (1995); Peacock, *J. Bone Miner. Res.*, 10:1294-1297 (1995)). Furthermore, the work thus far has not shed much light on the mechanism(s) whereby the genetic influences might exert their effect on bone mass.

While it is well known that peak bone mass is largely determined by genetic rather than environmental factors, studies to determine the gene loci (and ultimately the genes) linked to variation in bone mass are difficult and expensive. Study designs which utilize the power of linkage analysis, e.g., sib-pair or extended family, are generally more informative than simple association studies, although the latter do have value. However, genetic linkage studies involving bone mass are hampered by two major problems. The first problem is the phenotype, as discussed briefly above. Bone mass is a continuous, quantitative trait, and

establishing a discrete phenotype is difficult. Each anatomical site for measurement may be influenced by several genes, many of which may be different from site to site. The second problem is the age component of the phenotype. By the time an individual can be identified as having low bone mass, there is a high probability that their parents or other members of
5 prior generations will be deceased and therefore unavailable for study, and younger generations may not have even reached peak bone mass, making their phenotyping uncertain for genetic analysis.

Regardless, linkage analysis can be used to find the location of a gene causing a hereditary "disorder" and does not require any knowledge of the biochemical nature of the
10 disorder, i.e., a mutated protein that is believed to cause the disorder does not need to be known. Traditional approaches depend on assumptions concerning the disease process that might implicate a known protein as a candidate to be evaluated. The genetic localization approach using linkage analysis can be used to first find the general chromosomal region in which the defective gene is located and then to gradually reduce the size of the region in
15 order to determine the location of the specific mutated gene as precisely as possible. After the gene itself is discovered within the candidate region, the messenger RNA and the protein are identified and, along with the DNA, are checked for mutations.

The genetic localization approach has practical implications since the location of the disease can be used for prenatal diagnosis even before the altered gene that causes the disease
20 is found. Linkage analysis can enable families, even many of those that do not have a sick child, to know whether they are carriers of a disease gene and to evaluate the condition of an unborn child through molecular diagnosis. The transmission of a disease within families, then, can be used to find the defective gene. As used herein, reference to "high bone mass"

(HBM) is analogous to reference to a disease state, although from a practical standpoint high bone mass can actually help a subject avoid the disease known as osteoporosis.

Linkage analysis is possible because of the nature of inheritance of chromosomes from parents to offspring. During meiosis, the two parental homologues pair to guide their proper separation to daughter cells. While they are lined up and paired, the two homologues exchange pieces of the chromosomes, in an event called "crossing over" or "recombination." The resulting chromosomes are chimeric, that is, they contain parts that originate from both parental homologues. The closer together two sequences are on the chromosome, the less likely that a recombination event will occur between them, and the more closely linked they are. In a linkage analysis experiment, two positions on the chromosomes are followed from one generation to the next to determine the frequency of recombination between them. In a study of an inherited disease, one of the chromosomal positions is marked by the disease gene or its normal counterpart, i.e., the inheritance of the chromosomal region can be determined by examining whether the individual displays symptoms of the disorder or not. The other position is marked by a DNA sequence that shows natural variation in the population such that the two homologues can be distinguished based on the copy of the "marker" sequence that they possess. In every family, the inheritance of the genetic marker sequence is compared to the inheritance of the disease state. If, within a family carrying an autosomal dominant disorder such as high bone mass, every affected individual carries the same form of the marker and all the unaffected individuals carry at least one different form of the marker, there is a great probability that the disease gene and the marker are located close to each other. In this way, chromosomes may be systematically checked with known markers and compared to the disease state. The data obtained from the different families is combined, and analyzed together by a computer using statistical methods. The result is information

indicating the probability of linkage between the genetic marker and the disease allowing different distances between them. A positive result can mean that the disease is very close to the marker, while a negative result indicates that it is far away on that chromosome, or on an entirely different chromosome.

5 Linkage analysis is performed by typing all members of the affected family at a given marker locus and evaluating the co-inheritance of a particular disease state with the marker probe, thereby determining how often the two of them are co-inherited. The recombination frequency can be used as a measure of the genetic distance between two gene loci. A recombination frequency of 1% is equivalent to 1 map unit, or 1 centiMorgan (cM), which is
10 roughly equivalent to 1,000 kb of DNA. This relationship holds up to frequencies of about 20% or 20 cM.

The entire human genome is 3,300 cM long. In order to find an unknown disease gene within 5-10 cM of a marker locus, the whole human genome can be searched with roughly 330 informative marker loci spaced at approximately 10 cM intervals (Botstein *et al.*,
15 *Am. J. Hum. Genet.*, 32:314-331 (1980)). The reliability of linkage results is established by using a number of statistical methods. The method most commonly used for the analysis of linkage in humans is the LOD score method (Morton, *Prog. Clin. Biol. Res.*, 147:245-265 (1984), Morton *et al.*, *Am. J. Hum. Genet.*, 38:868-883 (1986)) which was incorporated into the computer program LIPED by Ott, *Am. J. Hum. Genet.*, 28:528-529 (1976). LOD scores
20 are the logarithm of the ratio of the likelihood that two loci are linked at a given distance to that they are not linked (>50 cM apart). The advantage of using logarithmic values is that they can be summed among families with the same disease. This becomes necessary given the relatively small size of human families.

By convention, a total LOD score greater than + 3.0 (that is, odds of linkage at the specified recombination frequency being 1000 times greater than odds of no linkage) is considered to be significant evidence for linkage at that particular recombination frequency. A total LOD score of less than - 2.0 (that is, odds of no linkage being 100 times greater than odds of linkage at the specified frequency) is considered to be strong evidence that the two loci under consideration are not linked at that particular recombination frequency. Until recently, most linkage analyses have been performed on the basis of two-point data, which is the relationship between the disorder under consideration and a particular genetic marker. However, as a result of the rapid advances in mapping the human genome over the last few years, and concomitant improvements in computer methodology, it has become feasible to carry out linkage analyses using multi-point data. Multi-point analysis provide a simultaneous analysis of linkage between the disease and several linked genetic markers, when the recombination distance among the markers is known.

Multi-point analysis is advantageous for two reasons. First, the informativeness of the pedigree is usually increased. Each pedigree has a certain amount of potential information, dependent on the number of parents heterozygous for the marker loci and the number of affected individuals in the family. However, few markers are sufficiently polymorphic as to be informative in all those individuals. If multiple markers are considered simultaneously, then the probability of an individual being heterozygous for at least one of the markers is greatly increased. Second, an indication of the position of the disease gene among the markers may be determined. This allows identification of flanking markers, and thus eventually allows isolation of a small region in which the disease gene resides. Lathrop *et al.*, *Proc. Natl. Acad. Sci. USA*, 81:3443-3446 (1984) have written the most widely used computer package, LINKAGE, for multi-point analysis.

There is a need in the art for identifying the gene associated with a high bone mass phenotype. The present invention is directed to this, as well as other, important ends.

SUMMARY OF THE INVENTION

The present invention describes the *Zmax1* gene and the *HBM* gene on chromosome 5 11q13.3 by genetic linkage and mutation analysis. The use of additional genetic markers linked to the genes has aided this discovery. By using linkage analysis and mutation analysis, persons predisposed to lipid associated disorders may be readily identified. Cloning methods using Bacterial Artificial Chromosomes have enabled the inventors to focus on the chromosome region of 11q13.3 and to accelerate the sequencing of the autosomal dominant 10 gene. In addition, the invention identifies the *Zmax1* gene and the *HBM* gene, and identifies the guanine-to-thymine polymorphism mutation at position 582 in the *Zmax1* gene that produces the *HBM* gene and the HBM phenotype as well as altered lipid levels.

The present invention identifies the *Zmax1* gene and the *HBM* gene, which can be used to determine if people are predisposed to abnormal lipid levels and, therefore, 15 susceptible to diseases mediated by lipids, including, for example, atherosclerosis, arteriosclerosis and associated conditions. Individuals with the *HBM* gene have lower LDL, triglyceride and VLDL levels and higher HDL levels. In other words, the *HBM* gene is a suppressor of atherosclerosis, arteriosclerosis and associated conditions. This *in vivo* observation is a strong evidence that treatment of normal individuals with the *HBM* gene or 20 protein, or fragments thereof, will ameliorate atherosclerosis, arteriosclerosis and conditions related thereto.

Moreover, such treatment will be indicated in the treatment of lipid-mediated diseases, particularly arteriosclerosis and conditions related thereto. For example, persons

predisposed to elevated lipid levels (i.e., diabetes, hypercholesteremia and other genetic diseases, obesity, male gender, and individuals who smoke) may be identified and/or treated by means of the invention. Moreover, the methods and compositions of the invention will be of use in the treatment or prevention of diabetic atherosclerotic disease, neurovascular conditions caused by plaque build-up (e.g., stroke), cardiovascular disease, poor circulation due to plaque build-up and associated poor wound healing.

In various embodiments, the present invention is directed to nucleic acids, proteins, vectors, and transformed hosts of HBM and Zmax1.

Additionally, the present invention is directed to applications of the above embodiments of the invention including, for example, gene therapy, pharmaceutical development, and diagnostic assays for bone development disorders. In preferred embodiments, the present invention is directed to methods for treating, diagnosing, preventing and screening for osteoporosis.

These and other aspects of the present invention are described in more detail below.

15 **BRIEF DESCRIPTION OF THE FIGURES**

Fig. 1 shows the pedigree of the individuals used in the genetic linkage studies. Under each individual is an ID number, the z-score for spinal BMD, and the allele calls for the critical markers on chromosome 11. Solid symbols represent "affected" individuals. Symbols containing "N" are "unaffected" individuals. DNA from 37 individuals was genotyped. Question marks denote unknown genotypes or individuals who were not genotyped.

Fig. 2 depicts the BAC/STS content physical map of the HBM region in 11q13.3. STS markers derived from genes, ESTs, microsatellites, random sequences, and BAC

endsequences are denoted above the long horizontal line. For markers that are present in GDB the same nomenclature has been used. Locus names (D11S####) are listed in parentheses after the primary name if available. STSs derived from BAC endsequences are listed with the BAC name first followed by L or R for the left and right end of the clone, respectively. The two large arrows indicate the genetic markers that define the HBM critical region. The horizontal lines below the STSs indicate BAC clones identified by PCR-based screening of a nine-fold coverage BAC library. Open circles indicate that the marker did not amplify the corresponding BAC library address during library screening. Clone names use the following convention: B for BAC, the plate, row and column address, followed by -H indicating the HBM project (i.e., B36F16-H).

Figs. 3A-3F show the genomic structure of *Zmax1* with flanking intron sequences. Translation is initiated by the underlined "ATG" in exon 1. The site of the polymorphism in the *HBM* gene is in exon 3 and is represented by the underlined "G," whereby this nucleotide is a "T" in the *HBM* gene. The 3' untranslated region of the mRNA is underlined within exon 23 (exon 1, SEQ ID NO:40; exon 2, SEQ ID NO:41; exon 3, SEQ ID NO:42; exon 4, SEQ ID NO:43; exon 5, SEQ ID NO:44; exon 6, SEQ ID NO:45; exon 7, SEQ ID NO:46; exon 8, SEQ ID NO:47; exon 9, SEQ ID NO:48; exon 10, SEQ ID NO:49; exon 11, SEQ ID NO:50; exon 12, SEQ ID NO:51; exon 13, SEQ ID NO:52; exon 14, SEQ ID NO:53; exon 15, SEQ ID NO:54; exon 16, SEQ ID NO:55; exon 17, SEQ ID NO:56; exon 18, SEQ ID NO:57; exon 19, SEQ ID NO:58; exon 20, SEQ ID NO:59; exon 21, SEQ ID NO:60; exon 22, SEQ ID NO:61; and exon 23; SEQ ID NO:62).

Fig. 4 shows the domain organization of *Zmax1*, including the YWTD spacers, the extracellular attachment site, the binding site for LDL and calcium, the cysteine-rich growth factor repeats, the transmembrane region, the ideal PEST region with the CK-II

phosphorylation site and the internalization domain. Fig. 4 also shows the site of the glycine to valine change that occurs in the HBM protein. The signal peptide is located at amino acids 1-22, the extracellular domain is located at amino acids 23-1385, the transmembrane segment is located at amino acids 1386-1413, and the cytoplasmic domain is located at amino acids

5 1414-1615.

Fig. 5 is a schematic illustration of the BAC contigs B527D12 and B200E21 in relation to the *HBM* gene.

Figs. 6A-6E are the nucleotide and amino acid sequences of the wild-type gene, *Zmax1*. The location for the base pair substitution at nucleotide 582, a guanine to thymine, is underlined. This allelic variant is the *HBM* gene. The *HBM* gene encodes for a protein with an amino acid substitution of glycine to valine at position 171. The 5' untranslated region (UTR) boundaries bases 1 to 70, and the 3' UTR boundaries bases 4916-5120.

10

Figs. 7A and 7B are northern blot analyses showing the expression of *Zmax1* in various tissues.

15 Fig. 8 is a PCR product analysis.

Fig. 9 is allele specific oligonucleotide detection of the *Zmax1* exon 3 mutation.

Fig. 10 is the cellular localization of mouse *Zmax1* by *in situ* hybridization at 100X magnification using sense and antisense probes.

Fig. 11 is the cellular localization of mouse *Zmax1* by *in situ* hybridization at 400X magnification using sense and antisense probes.

20

Fig. 12 is the cellular localization of mouse *Zmax1* by *in situ* hybridization of osteoblasts in the endosteum at 400X magnification using sense and antisense probes.

Fig. 13 shows antisense inhibition of *Zmax1* expression in MC-3T3 cells.

DETAILED DESCRIPTION OF THE INVENTION

To aid in the understanding of the specification and claims, the following definitions are provided.

"Gene" refers to a DNA sequence that encodes through its template or messenger RNA a sequence of amino acids characteristic of a specific peptide. The term "gene" includes intervening, non-coding regions, as well as regulatory regions, and can include 5' and 3' ends.

"Gene sequence" refers to a DNA molecule, including both a DNA molecule which contains a non-transcribed or non-translated sequence. The term is also intended to include any combination of gene(s), gene fragment(s), non-transcribed sequence(s) or non-translated sequence(s) which are present on the same DNA molecule.

The sequences of the present invention may be derived from a variety of sources including DNA, cDNA, synthetic DNA, synthetic RNA or combinations thereof. Such sequences may comprise genomic DNA which may or may not include naturally occurring introns. Moreover, such genomic DNA may be obtained in association with promoter regions or poly (A) sequences. The sequences, genomic DNA or cDNA may be obtained in any of several ways. Genomic DNA can be extracted and purified from suitable cells by means well known in the art. Alternatively, mRNA can be isolated from a cell and used to produce cDNA by reverse transcription or other means.

"cDNA" refers to complementary or copy DNA produced from an RNA template by the action of RNA-dependent DNA polymerase (reverse transcriptase). Thus, a "cDNA clone" means a duplex DNA sequence complementary to an RNA molecule of interest, carried in a cloning vector or PCR amplified. This term includes genes from which the intervening sequences have been removed.

"Recombinant DNA" means a molecule that has been recombined by *in vitro* splicing cDNA or a genomic DNA sequence.

"Cloning" refers to the use of *in vitro* recombination techniques to insert a particular gene or other DNA sequence into a vector molecule. In order to successfully clone a desired
5 gene, it is necessary to use methods for generating DNA fragments, for joining the fragments to vector molecules, for introducing the composite DNA molecule into a host cell in which it can replicate, and for selecting the clone having the target gene from amongst the recipient host cells.

"cDNA library" refers to a collection of recombinant DNA molecules containing
10 cDNA inserts which together comprise the entire genome of an organism. Such a cDNA library can be prepared by methods known to one skilled in the art and described by, for example, Cowell and Austin, "cDNA Library Protocols," Methods in Molecular Biology (1997). Generally, RNA is first isolated from the cells of an organism from whose genome it is desired to clone a particular gene.

15 "Cloning vehicle" refers to a plasmid or phage DNA or other DNA sequence which is able to replicate in a host cell. The cloning vehicle is characterized by one or more endonuclease recognition sites at which such DNA sequences may be cut in a determinable fashion without loss of an essential biological function of the DNA, which may contain a marker suitable for use in the identification of transformed cells.

20 "Expression control sequence" refers to a sequence of nucleotides that control or regulate expression of structural genes when operably linked to those genes. These include, for example, the lac systems, the trp system, major operator and promoter regions of the phage lambda, the control region of fd coat protein and other sequences known to control the expression of genes in prokaryotic or eukaryotic cells. Expression control sequences will

vary depending on whether the vector is designed to express the operably linked gene in a prokaryotic or eukaryotic host, and may contain transcriptional elements such as enhancer elements, termination sequences, tissue-specificity elements and/or translational initiation and termination sites.

5 "Expression vehicle" refers to a vehicle or vector similar to a cloning vehicle but which is capable of expressing a gene which has been cloned into it, after transformation into a host. The cloned gene is usually placed under the control of (i.e., operably linked to) an expression control sequence.

"Operator" refers to a DNA sequence capable of interacting with the specific
10 repressor, thereby controlling the transcription of adjacent gene(s).

"Promoter" refers to a DNA sequence that can be recognized by an RNA polymerase. The presence of such a sequence permits the RNA polymerase to bind and initiate transcription of operably linked gene sequences.

"Promoter region" is intended to include the promoter as well as other gene sequences
15 which may be necessary for the initiation of transcription. The presence of a promoter region is sufficient to cause the expression of an operably linked gene sequence.

"Operably linked" means that the promoter controls the initiation of expression of the gene. A promoter is operably linked to a sequence of proximal DNA if upon introduction into a host cell the promoter determines the transcription of the proximal DNA sequence(s)
20 into one or more species of RNA. A promoter is operably linked to a DNA sequence if the promoter is capable of initiating transcription of that DNA sequence.

"Prokaryote" refers to all organisms without a true nucleus, including bacteria.

"Eukaryote" refers to organisms and cells that have a true nucleus, including mammalian cells.

"Host" includes prokaryotes and eukaryotes, such as yeast and filamentous fungi, as well as plant and animal cells. The term includes an organism or cell that is the recipient of a replicable expression vehicle.

By "animal" is meant to include vertebrates. Preferred vertebrates include mammals and birds, but also include fish, reptiles and amphibians. Preferred mammals include:
5 humans, primates, rodents, canines, felines and livestock.

"Fragment" of a gene refers to any variant of the gene that possesses the biological activity of that gene.

"Variant" refers to a gene that is substantially similar in structure and biological
10 activity or immunological characteristics to either the entire gene or to a fragment of the gene. Provided that the two genes possess a similar activity, they are considered variant as that term is used herein even if the sequence of amino acid residues is not identical.

"Amplification of nucleic acids" refers to methods such as polymerase chain reaction (PCR), ligation amplification (or ligase chain reaction, LCR) and amplification methods
15 based on the use of Q-beta replicase. These methods are well known in the art and described, for example, in U.S. Patent Nos. 4,683,195 and 4,683,202. Reagents and hardware for conducting PCR are commercially available. Primers useful for amplifying sequences from the HBM region are preferably complementary to, and hybridize specifically to sequences in the HBM region or in regions that flank a target region therein. HBM sequences generated
20 by amplification may be sequenced directly. Alternatively, the amplified sequence(s) may be cloned prior to sequence analysis.

"Antibodies" may refer to polyclonal and/or monoclonal antibodies and fragments thereof, and immunologic binding equivalents thereof, that can bind to the HBM and Zmax1 proteins and fragments thereof or to nucleic acid sequences from the HBM or Zmax1 region,

particularly from the HBM locus or a portion thereof. The term antibody is used both to refer to a homogeneous molecular entity, or a mixture such as a serum product made up of a plurality of different molecular entities. Proteins may be prepared synthetically in a protein synthesizer and coupled to a carrier molecule and injected over several months into rabbits.

- 5 Rabbit sera is tested for immunoreactivity to the HBM protein or fragment. Monoclonal antibodies may be made by injecting mice with the proteins, or fragments thereof. Monoclonal antibodies will be screened by ELISA and tested for specific immunoreactivity with HBM protein or fragments thereof. Harlow *et al.*, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1988). These antibodies will be
- 10 useful in assays as well as pharmaceuticals. Antibodies can include antibody fragments (*e.g.*, scFv, Fab, F(ab')₂, *etc.*) as well as human antibodies, humanized antibodies and primatized antibodies.

"HBM" refers to high bone mass, but polymorphisms associated with *HBM* gene, which can also be involved in lipid modulation.

- 15 "HBM protein" refers to a protein that is identical to a *Zmax1* protein except that it contains an alteration of glycine 171 to valine. An HBM protein is defined for any organism that encodes a *Zmax1* true homologue. For example, a mouse HBM protein refers to the mouse *Zmax1* protein having the glycine 170 to valine substitution.

- "*HBM* gene" refers to the genomic DNA sequence found in individuals showing the
- 20 HBM characteristic or phenotype, where the sequence encodes the protein indicated by SEQ ID NO: 4. The *HBM* gene and the *Zmax1* gene are allelic. The protein encoded by the *HBM* gene has the property of causing elevated bone mass and also altering physiologic lipid levels, while the protein encoded by the *Zmax1* gene does not. The *HBM* gene and the *Zmax1* gene differ in that the *HBM* gene has a thymine at position 582, while the *Zmax1* gene has a

guanine at position 582. The *HBM* gene comprises the nucleic acid sequence shown as SEQ ID NO: 2. The *HBM* gene may also be referred to as an "HBM polymorphism."

"Normal," "wild-type," "unaffected" and "Zmax1" all refer to the genomic DNA sequence that encodes the protein indicated by SEQ ID NO: 3. The *Zmax1* gene has a
5 guanine at position 582. The *Zmax1* gene comprises the nucleic acid sequence shown as SEQ ID NO: 1. "Normal," "wild-type," "unaffected" and "Zmax1" also refer to allelic variants of the genomic sequence that encodes proteins that do not contribute to elevated bone mass. The *Zmax1* gene is common in the human population, while the *HBM* gene is rare.

"5YWT+EGF" refers to a repeat unit found in the Zmax1 protein, consisting of five
10 YWT repeats followed by an EGF repeat.

"Bone development" generally refers to any process involved in the change of bone over time, including, for example, normal development, changes that occur during disease states, and changes that occur during aging. "Bone development disorder" particularly refers to any disorders in bone development including, for example, changes that occur during
15 disease states and changes that occur during aging. Bone development may be progressive or cyclical in nature. Aspects of bone that may change during development include, for example, mineralization, formation of specific anatomical features, and relative or absolute numbers of various cell types.

"Bone modulation" or "modulation of bone formation" refers to the ability to affect
20 any of the physiological processes involved in bone remodeling, as will be appreciated by one skilled in the art, including, for example, bone resorption and appositional bone growth, by, inter alia, osteoclastic and osteoblastic activity, and may comprise some or all of bone formation and development as used herein.

"Normal bone density" refers to a bone density within two standard deviations or a \pm score of 0.

By "lipid regulation" or "lipid modulation" is meant the ability to alter by modulating the *HBM* or *Zmax1* genes, mRNA or protein encoded thereby the levels of a lipid. Altered
5 levels of lipid include very low density lipoproteins (VLDL), low density lipoproteins (LDL), high density lipoprotein (HDL) and triglycerides. The regulation or modulation can be an increase or decrease in the lipid level by an agent, which when administered to a subject modulates HBM or *Zmax1* activity. By "lipid metabolism" is meant the physiological cycle through which the various triglycerides and lipoproteins proceed. Agents of the invention
10 can also be said to modulate the metabolism of various lipids.

"Lipid" preferably includes very low density lipoproteins (VLDL), low density lipoproteins (LDL), intermediate density lipoprotein (IDL), high density lipoprotein (HDL) and triglycerides. Lipids can also include apolipoproteins, such as apolipoprotein A-1 (APO A-1), apolipoprotein B (APO B), apolipoprotein E (APO E) and lipoproteins such as lipoprotein
15 a (LIPOa).

By "lipid-mediated disease or condition" is meant to include arteriosclerosis and related conditions, hypercholesteremia, hyperlipidemia, atherosclerosis, and conditions or lifestyles associated with elevated lipid levels (e.g., diabetes mellitus, smoking and obesity) such as those discussed herein.

20 By "arteriosclerosis" is meant to include hypertrophy of the media and subintimal fibrosis with hyaline degeneration which can result in ectasia, aneurysm, increased systolic pressure, thrombus formation and embolism. Disorders associated with arteriosclerosis include, but are not limited to, nonatheromatous arteriosclerosis conditions such as: diabetes mellitus, chronic renal insufficiency, chronic vitamin D intoxication, pseudoxanthoma

elasticum, idiopathic arterial calcification in infancy, aortic valvular calcification in the elderly, and Werner's syndrome. Additional disorders associated with arteriosclerosis and atherosclerosis include: diabetes mellitus, hypertension, familial hypercholesterolemia, familial combined hyperlipidemia, familial dysbetalipoproteinemia, familial hypoalphalipoproteinemia, hypothyroidism, cholesterol ester storage disease, systemic lupus erythematosus and homocysteinemia.

By "atherosclerosis" is meant patchy intramural thickening of the subintima that encroaches on the arterial lumen and can cause obstruction. Atherosclerotic plaque consists of the accumulation of lipids, cells, connective tissue and glycosaminoglycans. It can cause the following conditions: stenosis, thrombosis, aneurysm, or embolus supervenes, as well as angina as well as the conditions listed above.

A "Zmax1 system" refers to a purified protein, cell extract, cell, animal, human or any other composition of matter in which Zmax1 is present in a normal or mutant form.

A "surrogate marker" refers to a diagnostic indication, symptom, sign or other feature that can be observed in a cell, tissue, human or animal that is correlated with the *HBM* gene or elevated bone mass or both, but that is easier to measure than bone density. The general concept of a surrogate marker is well accepted in diagnostic medicine.

The present invention encompasses the *Zmax1* gene and Zmax1 protein in the forms indicated by SEQ ID NOS: 1 and 3, respectively, and other closely related variants, as well as the adjacent chromosomal regions of Zmax1 necessary for its accurate expression. In a preferred embodiment, the present invention is directed to at least 15 contiguous nucleotides of the nucleic acid sequence of SEQ ID NO: 1.

The present invention also encompasses the *HBM* gene and HBM protein in the forms indicated by SEQ ID NO: 2 and 4, respectively, and other closely related variants, as well as

the adjacent chromosomal regions of the *HBM* gene necessary for its accurate expression. In a preferred embodiment, the present invention is directed to at least 15 contiguous nucleotides of the nucleic acid sequence of SEQ ID NO: 2. More preferably, the present invention is directed to at least 15 contiguous nucleotides of the nucleic acid sequence of
5 SEQ ID NO: 2, wherein one of the 15 contiguous nucleotides is the thymine at nucleotide 582.

The invention also relates to the nucleotide sequence of the *Zmax1* gene region, as well as the nucleotide sequence of the *HBM* gene region. More particularly, a preferred embodiment are the BAC clones containing segments of the *Zmax1* gene region B200E21-H
10 and B527D12-H. A preferred embodiment is the nucleotide sequence of the BAC clones consisting of SEQ ID NOS: 5-12.

The invention also concerns the use of the nucleotide sequence to identify DNA probes for the *Zmax1* gene and the *HBM* gene, PCR primers to amplify the *Zmax1* gene and the *HBM* gene, nucleotide polymorphisms in the *Zmax1* gene and the *HBM* gene, and
15 regulatory elements of the *Zmax1* gene and the *HBM* gene.

This invention describes the further localization of the chromosomal location of the *Zmax1* gene and *HBM* gene on chromosome 11q13.3 between genetic markers D11S987 and SNP_CONTIG033-6, as well as the DNA sequences of the *Zmax1* gene and the *HBM* gene. The chromosomal location was refined by the addition of more genetic markers to the
20 mapping panel used to map the gene, and by the extension of the pedigree to include more individuals. The pedigree extension was critical because the new individuals that have been genotyped harbor critical recombination events that narrow the region. To identify genes in the region on 11q13.3, a set of BAC clones containing this chromosomal region was identified. The BAC clones served as a template for genomic DNA sequencing, and also as a

reagent for identifying coding sequences by direct cDNA selection. Genomic sequencing and direct cDNA selection were used to characterize more than 1.5 million base pairs of DNA from 11q13.3. The *Zmax1* gene was identified within this region and the *HBM* gene was then discovered after mutational analysis of affected and unaffected individuals.

- 5 When a gene has been genetically localized to a specific chromosomal region, the genes in this region can be characterized at the molecular level by a series of steps that include: cloning of the entire region of DNA in a set of overlapping clones (physical mapping), characterization of genes encoded by these clones by a combination of direct cDNA selection, exon trapping and DNA sequencing (gene identification), and identification
- 10 of mutations in these genes by comparative DNA sequencing of affected and unaffected members of the HBM kindred (mutation analysis).

- Physical mapping is accomplished by screening libraries of human DNA cloned in vectors that are propagated in *E. coli* or *S. cerevisiae* using PCR assays designed to amplify unique molecular landmarks in the chromosomal region of interest. To generate a physical
- 15 map of the HBM candidate region, a library of human DNA cloned in Bacterial Artificial Chromosomes (BACs) was screened with a set of Sequence Tagged Site (STS) markers that had been previously mapped to chromosome 11q12-q13 by the efforts of the Human Genome Project.

- STSs are unique molecular landmarks in the human genome that can be assayed by
- 20 PCR. Through the combined efforts of the Human Genome Project, the location of thousands of STSs on the twenty-two autosomes and two sex chromosomes has been determined. For a positional cloning effort, the physical map is tied to the genetic map because the markers used for genetic mapping can also be used as STSs for physical mapping. By screening a BAC library with a combination of STSs derived from genetic markers, genes, and random

DNA fragments, a physical map comprised of overlapping clones representing all of the DNA in a chromosomal region of interest can be assembled.

BACs are cloning vectors for large (80 kilobase to 200 kilobase) segments of human or other DNA that are propagated in *E. coli*. To construct a physical map using BACs, a library of BAC clones is screened so that individual clones harboring the DNA sequence corresponding to a given STS or set of STSs are identified. Throughout most of the human genome, the STS markers are spaced approximately 20 to 50 kilobases apart, so that an individual BAC clone typically contains at least two STS markers. In addition, the BAC libraries that were screened contain enough cloned DNA to cover the human genome six times over. Therefore, an individual STS typically identifies more than one BAC clone. By screening a six-fold coverage BAC library with a series of STS markers spaced approximately 50 kilobases apart, a physical map consisting of a series of overlapping BAC clones, i.e. BAC contigs, can be assembled for any region of the human genome. This map is closely tied to the genetic map because many of the STS markers used to prepare the physical map are also genetic markers.

When constructing a physical map, it often happens that there are gaps in the STS map of the genome that result in the inability to identify BAC clones that are overlapping in a given location. Typically, the physical map is first constructed from a set of STSs that have been identified through the publicly available literature and World Wide Web resources. The initial map consists of several separate BAC contigs that are separated by gaps of unknown molecular distance. To identify BAC clones that fill these gaps, it is necessary to develop new STS markers from the ends of the clones on either side of the gap. This is done by sequencing the terminal 200 to 300 base pairs of the BACs flanking the gap, and developing a PCR assay to amplify a sequence of 100 or more base pairs. If the terminal sequences are

demonstrated to be unique within the human genome, then the new STS can be used to screen the BAC library to identify additional BACs that contain the DNA from the gap in the physical map. To assemble a BAC contig that covers a region the size of the HBM candidate region (2,000,000 or more base pairs), it is often necessary to develop new STS markers from
5 the ends of several clones.

After building a BAC contig, this set of overlapping clones serves as a template for identifying the genes encoded in the chromosomal region. Gene identification can be accomplished by many methods. Three methods are commonly used: (1) a set of BACs selected from the BAC contig to represent the entire chromosomal region can be sequenced,
10 and computational methods can be used to identify all of the genes, (2) the BACs from the BAC contig can be used as a reagent to clone cDNAs corresponding to the genes encoded in the region by a method termed direct cDNA selection, or (3) the BACs from the BAC contig can be used to identify coding sequences by selecting for specific DNA sequence motifs in a procedure called exon trapping. The present invention includes genes identified by the first
15 two methods.

To sequence the entire BAC contig representing the HBM candidate region, a set of BACs was chosen for subcloning into plasmid vectors and subsequent DNA sequencing of these subclones. Since the DNA cloned in the BACs represents genomic DNA, this sequencing is referred to as genomic sequencing to distinguish it from cDNA sequencing. To
20 initiate the genomic sequencing for a chromosomal region of interest, several non-overlapping BAC clones are chosen. DNA for each BAC clone is prepared, and the clones are sheared into random small fragments which are subsequently cloned into standard plasmid vectors such as pUC18. The plasmid clones are then grown to propagate the smaller fragments, and these are the templates for sequencing. To ensure adequate coverage and

sequence quality for the BAC DNA sequence, sufficient plasmid clones are sequenced to yield six-fold coverage of the BAC clone. For example, if the BAC is 100 kilobases long, then phagemids are sequenced to yield 600 kilobases of sequence. Since the BAC DNA was randomly sheared prior to cloning in the phagemid vector, the 600 kilobases of raw DNA
5 sequence can be assembled by computational methods into overlapping DNA sequences termed sequence contigs. For the purposes of initial gene identification by computational methods, six-fold coverage of each BAC is sufficient to yield ten to twenty sequence contigs of 1000 base pairs to 20,000 base pairs.

The sequencing strategy employed in this invention was to initially sequence "seed"
10 BACs from the BAC contig in the HBM candidate region. The sequence of the "seed" BACs was then used to identify minimally overlapping BACs from the contig, and these were subsequently sequenced. In this manner, the entire candidate region was sequenced, with several small sequence gaps left in each BAC. This sequence served as the template for computational gene identification. One method for computational gene identification is to
15 compare the sequence of BAC contig to publicly available databases of cDNA and genomic sequences, e.g. unigene, dbEST, genbank. These comparisons are typically done using the BLAST family of computer algorithms and programs (Altschul *et al.*, *J. Mol. Biol.*, 215:403-410 (1990)). The BAC sequence can also be translated into protein sequence, and the protein sequence can be used to search publicly available protein databases, using a version of
20 BLAST designed to analyze protein sequences (Altschul *et al.*, *Nucl. Acids Res.*, 25:3389-3402 (1997)). Another method is to use computer algorithms such as MZEF (Zhang, *Proc. Natl. Acad. Sci.*, 94:565-568 (1997)) and GRAIL (Uberbacher *et al.*, *Methods Enzymol.*, 266:259-281 (1996)), which predict the location of exons in the sequence based on the

presence of specific DNA sequence motifs that are common to all exons, as well as the presence of codon usage typical of human protein encoding sequences.

In addition to identifying genes by computational methods, genes were also identified by direct cDNA selection (Del Mastro *et al.*, *Genome Res.* 5(2):185-194 (1995)). In direct
5 cDNA selection, cDNA pools from tissues of interest are prepared, and the BACs from the candidate region are used in a liquid hybridization assay to capture the cDNAs which base pair to coding regions in the BAC. In the methods described herein, the cDNA pools were created from several different tissues by random priming the first strand cDNA from polyA RNA, synthesizing the second strand cDNA by standard methods, and adding linkers to the
10 ends of the cDNA fragments. The linkers are used to amplify the cDNA pools. The BAC clones are used as a template for *in vitro* DNA synthesis to create a biotin labelled copy of the BAC DNA. The biotin labelled copy of the BAC DNA is then denatured and incubated with an excess of the PCR amplified, linkered cDNA pools which have also been denatured. The BAC DNA and cDNA are allowed to anneal in solution, and heteroduplexes between the
15 BAC and the cDNA are isolated using streptavidin coated magnetic beads. The cDNAs that are captured by the BAC are then amplified using primers complimentary to the linker sequences, and the hybridization/selection process is repeated for a second round. After two rounds of direct cDNA selection, the cDNA fragments are cloned, and a library of these direct selected fragments is created.

20 The cDNA clones isolated by direct selection are analyzed by two methods. Since a pool of BACs from the HBM candidate region is used to provide the genomic DNA sequence, the cDNAs must be mapped to individual BACs. This is accomplished by arraying the BACs in microtiter dishes, and replicating their DNA in high density grids. Individual cDNA clones are then hybridized to the grid to confirm that they have sequence identity to an

individual BAC from the set used for direct selection, and to determine the specific identity of that BAC. cDNA clones that are confirmed to correspond to individual BACs are sequenced. To determine whether the cDNA clones isolated by direct selection share sequence identity or similarity to previously identified genes, the DNA and protein coding
5 sequences are compared to publicly available databases using the BLAST family of programs.

The combination of genomic DNA sequence and cDNA sequence provided by BAC sequencing and by direct cDNA selection yields an initial list of putative genes in the region. The genes in the region were all candidates for the HBM locus. To further characterize each
10 gene, Northern blots were performed to determine the size of the transcript corresponding to each gene, and to determine which putative exons were transcribed together to make an individual gene. For Northern blot analysis of each gene, probes were prepared from direct selected cDNA clones or by PCR amplifying specific fragments from genomic DNA or from the BAC encoding the putative gene of interest. The Northern blots gave information on the
15 size of the transcript and the tissues in which it was expressed. For transcripts which were not highly expressed, it was sometimes necessary to perform a reverse transcription PCR assay using RNA from the tissues of interest as a template for the reaction.

Gene identification by computational methods and by direct cDNA selection provides unique information about the genes in a region of a chromosome. When genes are identified,
20 then it is possible to examine different individuals for mutations in each gene.

I. Phenotyping using DXA Measurements

Spinal bone mineral content (BMC) and bone mineral density (BMD) measurements performed at Creighton University (Omaha, Nebraska) were made by DXA using a Norland

Instruments densitometer (Norland XR2600 Densitometer, Dual Energy X-ray Absorptiometry, DXA). Spinal BMC and BMD at other locations used the machinery available. There are estimated to be 800 DXA machines currently operating in the U.S. Most larger cities have offices or imaging centers which have DXA capabilities, usually a Lunar or Hologic machine. Each location that provided spine BMC and BMD data included copies of the printouts from their machines to provide verification that the regions of interest for measurement of BMD have been chosen appropriately. Complete clinical histories and skeletal radiographs were obtained.

The HBM phenotype is defined by the following criteria: very high spinal BMD; a clinical history devoid of any known high bone mass syndrome; and skeletal radiographs showing a normal shape of the appendicular skeleton.

II. Genotyping of Microsatellite Markers

To narrow the genetic interval to a region smaller than that originally reported by Johnson *et al.*, *Am. J. Hum. Genet.*, 60:1326-1332 (1997), additional microsatellite markers on chromosome 11q12-13 were typed. The new markers included: D11S4191, D11S1883, D11S1785, D11S4113, D11S4136, D11S4139, (Dib, *et al.*, *Nature*, 380:152-154 (1996), FGF3 (Polymeropolous, *et al.*, *Nucl. Acid Res.*, 18:7468 (1990)), as well as GTC_HBM_Marker_1, GTC_HBM_Marker_2, GTC_HBM_Marker_3, GTC_HBM_Marker_4, GTC_HBM_Marker_5, GTC_HBM_Marker_6, and GTC_HBM_Marker_7 (See Fig. 2).

Blood (20 ml) was drawn into lavender cap (EDTA containing) tubes by a certified phlebotomist. The blood was stored refrigerated until DNA extraction. DNA has been extracted from blood stored for up to 7 days in the refrigerator without reduction in the

quality or quantity of yield. For those subjects that have blood drawn at distant sites, a shipping protocol was successfully used on more than a dozen occasions. Blood samples were shipped by overnight express in a styrofoam container with freezer packs to provide cooling. Lavender cap tubes were placed on individual plastic shipping tubes and then into
5 "zip-lock" biohazard bags. When the samples arrived the next day, they were immediately processed to extract DNA.

The DNA extraction procedure used a kit purchased from Gentra Systems, Inc. (Minneapolis, Minnesota). Briefly, the procedure involved adding 3 volumes of a red blood cell lysis buffer to the whole blood. After incubations for 10 minutes at room temperature,
10 the solution was centrifuged in a Beckman tabletop centrifuge at 2,000 X g for 10 minutes. The white blood cell pellet was resuspended in Cell Lysis Buffer. Once the pellet was completely resuspended and free of cell clumps, the solution was digested with RNase A for 15 minutes at 37°C. Proteins were precipitated by addition of the provided Protein Precipitation Solution and removed by centrifugation. The DNA was precipitated out of the
15 supernatant by addition of isopropanol. This method was simple and fast, requiring only 1-2 hours, and allowed for the processing of dozens of samples simultaneously. The yield of DNA was routinely >8 mg for a 20 ml sample of whole blood and had a MW of >50 kb. DNA was archived by storing coded 50 µg aliquots at -80°C as an ethanol precipitate.

DNA was genotyped using one fluorescently labeled oligonucleotide primer and one
20 unlabeled oligonucleotide primer. Labeled and unlabeled oligonucleotides were obtained from Integrated DNA Technologies, Inc. (Coralville, Iowa). All other reagents for microsatellite genotyping were purchased from Perkin Elmer-Applied Biosystems, Inc. ("PE-ABI") (Norwalk, Connecticut). Individual PCR reactions were performed for each marker, as described by PE-ABI using AmpliTaq DNA Polymerase. The reactions were added to 3.5 µl

of loading buffer containing deionized formamide, blue dextran and TAMRA 350 size standards (PE-ABI). After heating at 95°C for 5 minutes to denature the DNA, the samples were loaded and electrophoresed as described in the operator's manual for the Model 377 DNA Sequencer (PE-ABI, Foster City, California). After gel electrophoresis, the data was
5 analyzed using PE-ABI GENESCAN™ and GENOTYPER™ software. First, within the GENESCAN™ software, the lane tracking was manually optimized prior to the first step of analysis. After the gel lane data was extracted, the standard curve profiles of each lane were examined and verified for linearity and size calling. Lanes, which had problems with either of these parameters, were re-tracked and verified. Once all lanes were tracked and the size
10 standards were correctly identified, the data were imported into GENOTYPER™ for allele identification. To expedite allele calling (binning), the program Linkage Designer from the Internet web-site of Dr. Guy Van Camp (<http://alt.www.uia.ac.be/u/dnalab/ld.html>) was used. This program greatly facilitates the importing of data generated by GENOTYPER™ into the pedigree drawing program Cyrillic (Version 2.0, Cherwell Scientific Publishing Limited,
15 Oxford, Great Britain) and subsequent linkage analysis using the program LINKAGE (Lathrop *et al.*, *Am. J. Hum. Genet.*, 37:482-498 (1985)).

III. Linkage Analysis

Fig. 1 demonstrates the pedigree of the individuals used in the genetic linkage studies for this invention. Specifically, two-point linkage analysis was performed using the MLINK
20 and LINKMAP components of the program LINKAGE (Lathrop *et al.*, *Am. J. Hum. Genet.*, 37:482-498 (1985)). Pedigree/marker data was exported from Cyrillic as a pre-file into the Makeped program and converted into a suitable ped-file for linkage analysis.

The original linkage analysis was performed using three models: (i) an autosomal dominant, fully penetrant model, (ii) an autosomal dominant model with reduced penetrance, and (iii) a quantitative trait model. The HBM locus was mapped to chromosome 11q12-13 by analyzing DNA for linked markers from 22 members of a large, extended kindred. A highly automated technology was used with a panel of 345 fluorescent markers which spanned the 22 autosomes at a spacing interval ranging from 6-22 cM. Only markers from this region of chromosome 11 showed evidence of linkage (LOD score ~3.0). The highest LOD score (5.74) obtained by two-point and multipoint analysis was D11S987 (map position 55 in Fig. 2). The 95% confidence interval placed the HBM locus between markers D11S905 and D11S937 (map position 41-71 in Fig. 2). Haplotype analysis also places the *Zmax1* gene in this same region. Further descriptions of the markers D11S987, D11S905, and D11S937 can be found in Gyapay *et al.*, *Nature Genetics*, Vol. 7, (1994).

In this invention, the inventors report the narrowing of the HBM interval to the region between markers D11S987 and GTC_HBM_Marker_5. These two markers lie between the delimiting markers from the original analysis (D11S11S905 and D11S937) and are approximately 3 cM from one another. The narrowing of the interval was accomplished using genotypic data from the markers D11S4191, D11S1883, D11S1785, D11S4113, D11S4136, D11S4139, (Dib *et al.*, *Nature*, 380:152-154 (1996)), FGF3 (Polymeropoulos *et al.*, *Nucl. Acid Res.*, 18:7468 (1990)) (information about the genetic markers can be found at the internet site of the Genome Database, <http://gdbwww.gdb.org/>), as well as the markers GTC_HBM_Marker_1, GTC_HBM_Marker_2, GTC_HBM_Marker_3, GTC_HBM_Marker_4, GTC_HBM_Marker_5, GTC_HBM_Marker_6, and GTC_HBM_Marker_7.

As shown in Fig. 1, haplotype analysis with the above genetic markers identifies recombination events (crossovers) in individuals 9019 and 9020 that significantly refine the interval of chromosome 11 to which the *Zmax1* gene is localized. Individual 9019 is an HBM-affected individual that inherits a portion of chromosome 11 from the maternal chromosome with the *HBM* gene, and a portion from the chromosome 11 homologue. The portion inherited from the *HBM* gene-carrying chromosome includes markers D11S935, D11S1313, GTC_HBM_Marker_4, D11S987, D11S1296, GTC_HBM_Marker_6, GTC_HBM_Marker_2, D11S970, GTC_HBM_Marker_3, D11S4113, GTC_HBM_Marker_1, GTC_HBM_Marker_7 and GTC_HBM_Marker_5. The portion from D11S4136 and continuing in the telomeric direction is derived from the non-HBM chromosome. This data places the *Zmax1* gene in a location centromeric to the marker GTC_HBM_Marker_5. Individual 9020 is an unaffected individual who also exhibits a critical recombination event. This individual inherits a recombinant paternal chromosome 11 that includes markers D11S935, D11S1313, GTC_HBM_Marker_4, D11S987, D11S1296 and GTC_HBM_Marker_6 from her father's (individual 0115) chromosome 11 homologue that carries the *HBM* gene, and markers GTC_HBM_Marker_2, D11S970, GTC_HBM_Marker_3, GTC_HBM_Marker_1, GTC_HBM_Marker_7, GTC_HBM_Marker_5, D11S4136, D11S4139, D11S1314, and D11S937 from her father's chromosome 11 that does not carry the *HBM* gene. Marker D11S4113 is uninformative due to its homozygous nature in individual 0115. This recombination event places the centromeric boundary of the HBM region between markers D11S1296 and D11S987.

Two-point linkage analysis was also used to confirm the location of the *Zmax1* gene on chromosome 11. The linkage results for two point linkage analysis under a model of full penetrance are presented in Table 1 below. This table lists the genetic markers in the first

column and the recombination fractions across the top of the table. Each cell of the column shows the LOD score for an individual marker tested for linkage to the *Zmax1* gene at the recombination fraction shown in the first row. For example, the peak LOD score of 7.66 occurs at marker D11S970, which is within the interval defined by haplotype analysis.

TABLE 1

Marker	0.0	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4
D11S935	- infinity	0.39	0.49	0.47	0.41	0.33	0.25	0.17	0.10
D11S1313	- infinity	2.64	2.86	2.80	2.59	2.30	1.93	1.49	1.00
D11S987	- infinity	5.49	5.18	4.70	4.13	3.49	2.79	2.03	1.26
D11S4113	4.35	3.99	3.62	3.24	2.83	2.40	1.94	1.46	0.97
D11S1337	2.29	2.06	1.81	1.55	1.27	0.99	0.70	0.42	0.18
D11S970	7.66	6.99	6.29	5.56	4.79	3.99	3.15	2.30	1.44
D11S4136	6.34	5.79	5.22	4.61	3.98	3.30	2.59	1.85	1.11
D11S4139	6.80	6.28	5.73	5.13	4.50	3.84	3.13	2.38	1.59
FGF3	0.59	3.23	3.15	2.91	2.61	2.25	1.84	1.40	0.92
D11S1314	6.96	6.49	5.94	5.34	4.69	4.01	3.27	2.49	1.67
D11S937	-infinity	4.98	4.86	4.52	4.06	3.51	2.88	2.20	1.47

A single nucleotide polymorphism (SNP) further defines the HBM region. This SNP is termed SNP_Contig033-6 and is located 25 kb centromeric to the genetic marker GTC_HBM_Marker_5. This SNP is telomeric to the genetic marker GTC_HBM_Marker_7. SNP_Contig033-6 is present in HBM-affected individual 0113. However, the HBM-affected individual 9019, who is the son of 0113, does not carry this SNP. Therefore, this indicates that the crossover is centromeric to this SNP. The primer sequence for the genetic markers GTC_HBM_Marker_5 and GTC_HBM_Marker_7 is shown in Table 2 below.

TABLE 2

Marker	Primer (Forward)	Primer (Reverse)
GTC_HBM_Marker_5	TTTGGGTACACAATTCAGTCG	AAAACTGTGGGTGCTTCTGG
GTC_HBM_Marker_7	GTGATTGAGCCAATCCTGAGA	TGAGCCAAATAAACCCCTTCT

5 The kindred described have several features of great interest, the most important being that their bones, while very dense, have an absolutely normal shape. The outer dimensions of the skeletons of the HBM-affected individuals are normal, and, while medullary cavities are present, there is no interference with hematopoiesis. The HBM-affected members seem to be resistant to fracture, and there are no neurologic symptoms, and no symptoms of impairment

10 of any organ or system function in the members examined. HBM-affected members of the kindred live to advanced age without undue illness or disability. Furthermore, the HBM phenotype matches no other bone disorders such as osteoporosis, osteoporosis pseudoglioma, Engelmann's disease, Ribbing's disease, hyperphosphatasemia, Van Buchem's disease, melorheostosis, osteopetrosis, pycnodysostosis, sclerostenosis, osteopoikilosis, acromegaly,

15 Paget's disease, fibrous dysplasia, tubular stenosis, osteogenesis imperfecta, hypoparathyroidism, pseudohypoparathyroidism, pseudopseudohypoparathyroidism, primary and secondary hyperparathyroidism and associated syndromes, hypercalciuria, medullary carcinoma of the thyroid gland, osteomalacia and other diseases. Clearly, the HBM locus in this family has a very powerful and substantial role in regulating bone density, and its

20 identification is an important step in understanding the pathway(s) that regulate bone density and the pathogenesis of diseases such as osteoporosis.

In addition, older individuals carrying the *HBM* gene, and therefore expression of the HBM protein, do not show loss of bone mass characteristic of normal individuals. Moreover, individuals carrying the *HBM* gene have lower triglycerides, VLDLs, and LDLs and/or

increased HDLs. In other words, the *HBM* gene is a suppressor of osteoporosis and may lessen cardiovascular risk arteriosclerotic and/or atherosclerotic associated conditions. In essence, individuals carrying the *HBM* gene are dosed with the HBM protein, and, as a result, lower levels of detrimental lipids (*e.g.*, VLDL, LDL and triglycerides). This *in vivo* observation is strong evidence that treatment of normal individuals with the *HBM* gene or protein, or a fragment thereof, will ameliorate osteoporosis and arterio- or atherosclerotic conditions or diseases.

IV. Physical Mapping

To provide reagents for the cloning and characterization of the HBM locus, the genetic mapping data described above were used to construct a physical map of the region containing *Zmax1* on chromosome 11q13.3. The physical map consists of an ordered set of molecular landmarks, and a set of BAC clones that contain the *Zmax1* gene region from chromosome 11q13.3.

Various publicly available mapping resources were utilized to identify existing STS markers (Olson *et al.*, *Science*, 245:1434-1435 (1989)) in the HBM region. Resources included the GDB, the Whitehead Institute Genome Center, dbSTS and dbEST (NCBI), 11db, the University of Texas Southwestern GESTEC, the Stanford Human Genome Center, and several literature references (Courseaux *et al.*, *Genomics*, 40:13-23 (1997), Courseaux *et al.*, *Genomics*, 37:354-365 (1996), Guru *et al.*, *Genomics*, 42:436-445 (1997), Hosoda *et al.*, *Genes Cells*, 2:345-357 (1997), James *et al.*, *Nat. Genet.*, 8:70-76 (1994), Kitamura *et al.*, *DNA Research*, 4:281-289 (1997), Lemmens *et al.*, *Genomics*, 44:94-100 (1997), Smith *et al.*, *Genome Res.*, 7:835-842 (1997)). Maps were integrated manually to identify markers mapping to the region containing *Zmax1*.

Primers for existing STSs were obtained from the GDB or literature references are listed in Table 3 below. Thus, Table 3 shows the STS markers used to prepare the physical map of the *Zmax1* gene region.

TABLE 3: HBM STS Table

STS Name	Locid Name	Type	Gene Access.	Size (kb)	Forward Primer	Reverse Primer	Gene Name
ACTN3		Gene	GDB:197568	0.184	CTGGACCTAGTGGCCTCTC	TTGAGAGCACTTGGCTGG	Actinin, alpha 3 - skeletal muscle
PC-BIPC-Y		Gene	GDB:197084	0.125	CTCAGTCCGATGAGATGGA	CAAGATCACTGATCTCCAGG	Pyruvate Carboxylase
ADRBK1	D11S2161	Gene	GDB:4580179	0.322	GTTCAGGAGACTCAGATC	TTCTGCAAGTGTCTGTGAG	Adenosine Receptor (A2) Gene
PSANK3		Gene	GDB:197568	0.259	GAGAAAGAAATAGGGGACC	GGCCTCTGCTGACTTCAGG	Beta-adrenergic receptor kinase
PP11/2/PP1(2/2)		Gene	GDB:197568	0.208	GAGTACGGGCTGACGTGGCT	TTCTTTGAAGGACATGAGA	sim. to Human endogenous retrovirus mRNA long terminal repeat
GSTP1.PCR1		Gene	GDB:270056	0.19	AGCTGGCCACAGCTGAGATAC	ATACCAAGAGTGCATGTTCCCGT	Protein phosphatase 1, catalytic subunit, alpha isoform
NDUFV1		Gene	GDB:197568	0.521	CATGTGCTCCAGCTCATCAT	CAAGATCTGTAGTCTCTG	Glutathione S-transferase p1
PSANK2		Gene	GDB:197568	0.157	CAGAGAACTCAAGGGACTTG	ATCTCTCACTCCACACT	NADH dehydrogenase (ubiquinone) flavoprotein 1 (51kD)
PSANK1		EST	GDB:314521	0.3	CAAGGCTAAAGACGAAAA	TCAGAGGATTCATCATTTT	Adenylate dehydrogenase 2 (ALDH8)
UT5620	D11S1917	MSAT	GDB:314521	0.211	AAGTGGAGGCTGCAAGAG	GGCCTGTGCTCTTCAGTA	Human ribosomal protein L37 (PSANK1) pseudogene.
AFM289va9	D11S1937	MSAT	GDB:199905	0.287	AAGGTGTGAGGATCACTGG	AGCTCATGGGGCTTATT	Preprogalanin (GAL1)
GALN		Gene	GDB:177850	0.322	GTCTCTCCGAGTGTATCAAC	ATGGCAGGAGCTTAGAACA	Protein D1
PM551	D11S97	EST	GDB:177850	0.205	GATCAGCGAAGTCTCTCTGGCTC	TCACATGAGGACTGTGGAGCG	Cyclin D1
BCL1(1)/BCL1(2)		Gene	GDB:4580141	0.248	GATCAGCGAAGTCTCTCTGGCTC	TCACATGAGGACTGTGGAGCG	Fibroblast growth factor 4
FGF3.PCR1		Gene	GDB:4580141	0.248	GATCAGCGAAGTCTCTCTGGCTC	TCACATGAGGACTGTGGAGCG	Fibroblast growth factor 3
AFM164ZF12	D11S913	MSAT	GDB:188151	0.22	CATTGGGAAATCCAGAGA	TAGGTCTCTATTTTGTGCTTG	Choline Kinase
AFM190YD5		MSAT	GDB:122329	0.275	GACATACCATGACACTATAAGAG	CAACCCATACCAAGGATAAG	
SHGC-18285	D11S4689	STS	GDB:740600	0.147	GACACAGAGGGTAACTTGGC	TGAGGACACAGATCTGATGGG	
SHGC-3084	D11S4640	STS	GDB:740102	0.167	GAGGTGCTCTCTTAATCTTTG	GACTATATGTAGTAGTAGGAG	
SHGC-14407	D11S4654	STS	GDB:740518	0.158	CTCTGTAACCCCACTGCC	TCCTGCTCTCAAGTCTCTGG	
SHGC-10945	D11S4327	Gene	GDB:674522	0.311	ACTCATGACCTCATCACTG	TGCTGTTGCCCTCATCTGAC	
SS15	D11S103	STS	GDB:198260	0.165	GTGACAGGCAATAGCTGAGG	TGCTGCTCTCTGCTGCTGAG	
AFM147X10	D11S1888	MSAT	GDB:307895	0.183	AGCTGGATCTCTCAGAAATG	CAAGAGGCTGTAGAGAGTG	
AFM131YE5	D11S887	MSAT	GDB:311922	0.237	GAGTCCAGTCTGGCAATAAAGC	GGTGGGACATGACCTCTAAG	
AFM359x9a9	D11S4178	MSAT	GDB:608115	0.237	GAGTCCAGTCTGGCAATAAAGC	GGTGGGACATGACCTCTAAG	
AFM272v05	D11S4113	MSAT	GDB:458164	0.218	ACTCAGCGGTGTATCGC	CTTGAAGCCGATCTTTCG	
SGC31923		EST	GDB:458164	0.15	TATTTGCAAGCTTGAGACTTCT	ATCAGTGTGCTTTGTGCTC	
WI-7741	D11S4364	Gene	GDB:677552	0.125	ACTTATTTGACGCTGGGCG	CCCACTGGCTGTGTTTATG	Transformation-sensitive protein IEF SSP 3521
SGC35223		EST	GDB:4582598	0.324	GAGCAGGGGAGAGAGAGGC	CCCACTGGCTGTGTTTATG	ZNF 162 - splicing factor 1
WI-16754		EST	GDB:4582598	0.15	AGCCACTTATTTATTTTATGTC	CCCACTGGCTGTGTTTATG	
WI-6315	D11S4418	EST	GDB:678804	0.224	ATGCTTTTGCATGATTTAATTT	TACTGTCTGATAGTAGTGTGGC	
WI-16915		EST	GDB:4584055	0.125	CTGCTTCTCTGCTGCTGCTG	TCGCCCAAGAAATGTAAAGG	
SGC3080B		EST	GDB:4584055	0.128	TCAGAGCAGCACTGTTTAAACA	ATCCACGCGCCAGGAT	Mitogen Inducible gene (MIG-2)
WI-17663		EST	GDB:4583346	0.128	CAAGCGGGTGTATTTAACA	CTGCTGAAAGTCTAGAGCC	
SGC31567		Gene	GDB:1222237	0.199	GCATATAGAACTTATTTGCCG	GATGCCAGGACCATGAGC	Human tel interactiv protein (TIP60)
SGC30556		Gene	GDB:4578432	0.207	GTACCAACACACAGCAGGC	CTCTGAAGCAGGAGGAGAG	Calcium activated neutral protease large subunit, muCAHP, calpain
SGC33927		EST	GDB:4584037	0.15	GTGCTTGTGACTCAGGCTGTC	TTTCCCTCAACAATCACTACTCC	
WI-8671		EST	GDB:4582382	0.15	GGTGGGGATATAGAGGTCA	TACTGCGCCAGAGAGCTAGG	
WI-12334		EST	GDB:1222235	0.124	TGGTTTAAACCTTAAAGAGAA	AGCTTGAGATGAGAGGCC	
WI-16402		EST	GDB:4581874	0.113	GGTAAAGAAACATTTGAGAGAT	TGTTGATCTACCTGCTTCCG	
WI-16671		EST	GDB:4584947	0.131	TTGAAACCAATTTTTCACCG	TGTTGATCTACCTGCTTCCG	
WI-12656		EST	GDB:458106	0.209	TTGAAACCAATTTTTCACCG	TCTGCGGCTGTGGATTT	HLA-B
SGC33767		EST	GDB:458106	0.15	CTTATTTGAAACATTTGAGTGA	TGTTGATCTACCTGCTTCCG	HLA-B
AFM343YB5		MSAT	GDB:1222332	0.181	AAACCAACACACCA	TTTCCCTCAACAATCACTACTCC	
SGC33744		EST	GDB:4581592	0.15	CTTGTGTAGAGCAAGGCTCA	CCCTGGAAAGGTGAATGCT	
SGC32272		EST	GDB:4583084	0.135	GACGAGGTGATTCAGGCG	TATCTGTGTGCTGCTTCAATGT	
SGC34148		EST	GDB:4583084	0.135	GACGAGGTGATTCAGGCG	ACTGAAGACTCTGCTCCT	
WI-18546		EST	GDB:4574588	0.133	TATGTAAGCAATAGTGAACCC	CACCTCTCCCACTTTGTC	Human 1.1 kb mRNA upregulated in retinoic acid treated HL-60 neutrophilic cells
SGC31103		EST	GDB:4587265	0.1	CTATGCCAGAGATGACAGG	TGGCAATGAGGCTATGTGCG	Human pyruvate carboxylase precursor
SGC30028		Gene	GDB:4585055	0.128	GCAGGCTTATTTAGTAAAGTCC	CACGTGAGACTCAAGTGTGG	
WI-2075	D11S4407	STS	GDB:677554	0.125	CATCCCAACATCACTCAGT	GGGGACTAGCTTACAGATTGA	LAR-interacting protein 1b
SGC36955		Gene	GDB:4577182	0.223	AGATCACTTTCGACAGCTGG	TGAAGGATATTTAGAGCTTGA	
GCT16807	D11S4270	STS	GDB:628245	0.197	GAGGTTTTCCTCCGATC	TGAGGTTGGAGAGATCAT	
WI-6504	D11S3974	EST	GDB:588142	0.174	CTTCTATAGCCACACCGG	CAGCTACTGTGAGATGCA	
SGC31049		EST	GDB:4580093	0.15	TCCTTACTGCTTACCACTTCC	CAACAGTGCATCGGTATCG	

TABLE 3: HBM STS Table

TIGR-AD02J17	EST	DBP:1222193	0.199	AGATGAGCAGCAGATAG	CATTCCACATGGATAGAC	NDUFV1 amplixin (EMS1)
W-5998	EST	DBP:459683	0.1	CATACATGAGGTGCTACAGG	GCATTTCTCATCTCTCTGC	Nuclear mitotic apparatus protein 1, NUMA
W-16987	EST	DBP:4575848	0.15	ITACAGCCCAAGGTGCTTC	AGGTGTGTGTCAGAGTGA	
SGC31912	EST	DBP:4597868	0.101	CACGTGATGCTATGCTGTGAGG	TTGATTTGTGCTCCGAAA	
W-13500	EST	DBP:4577893	0.15	CCGACACCTCCACCTTTTATTT	CCAGTCACCTTACTAGCTCTTGG	
CHLC.GAAT1B01.P7933	MSAT	DBP:684255	0.103	AGGACACAGCTGCTCATATG	ACCAGGATTCGCTCAAAAG	
SGC35519	Gene	DBP:4577180	0.134	GATGGGTGCTGCTGCTGCTCA	ACATTTATTTGACATGCAACC	LAR-interacting protein 1a mRNA
W-11974	EST	DBP:1222255	0.108	AGCATCTTAACTGTGTCAGGCA	ATGTGCTGGGCTGGAAG	Caroline palmitoyl transferase 1
W-15244	Gene	DBP:4574740	0.108	TCACATGCAAAATCGGCA	CTGGCTGTGTGGTGTGCG	Beta-adrenergic receptor kinase 1, ADRB1
W-17466	EST	DBP:4583336	0.131	TTGTTTATTTCTGATACAAAGCCA	GACCTCTGTGACACACG	FGF4
W-9159	EST	DBP:678144	0.111	CCACCAATTAATTAATGCTGTGG	GTAGATCTCCAGTGTGACCC	
W-4232	STS	DBP:1222250	0.175	CTATATGAGGCTGGACCA	ACTGCTCATGTGAGTACCCG	
SHGC-4187	EST	DBP:4566789	0.161	CATGTGTGCTGCTGCTTCTT	CAGCATCTGACGCTTACG	Human DNA helicase gene (SMBP2)
W-14303	EST	DBP:4576938	0.15	CTGCTATTAATGAGAACTACAG	TCCTCTGGGAGTCAGAGTC	
W-16597	EST	DBP:4585666	0.13	CAGGGCAGTGAATACACTACG	AAGGATCAAGCAGGCACTTG	
RC29S1CATFORRC29S1CAT	MSAT	DBP:191084	0.15	ACACATCTCTGCTGCTCC	TGAACCTGGAGGCGAG	
UT978	MSAT	DBP:198525	0.392	CATTCCTGCTGCTGCTGCT	GTGCTGGGATACAGGTGT	
1261/1282	EST	DBP:335216	0.07	GCAGAGGCTCTGCTAGCC	CCATGCTAGAGAGCAGAC	
D11S468	STS	DBP:179349	0.098	AGTGTGGGCGAGGACCTCTG	CAGACAGATACCTGGGCTC	
D11S668	STS	DBP:4572653	0.143	TCCCTCATCTGCTGCTGCT	AGCCCCCTGGGGATAATC	
RH18048	Gene	DBP:4572653	0.188	GATGCTTACTTACACGCG	AGGATCTATCTGGGCTATG	Aldohyde dehydrogenase (ALDH8)
IGHMBP2	Gene	DBP:4590087	0.699	TGCAGACCATGCTCCGCT	GAGAGGCGCGGAGGCTCTG	Human DNA helicase gene (SMBP2)
KRMI	Gene	DBP:4590244	0.277	CTCATCAACACCATGTTGAGGCT	GGGTGTGAGCTGCTGCTGGAAG	Nuclear mitotic apparatus protein 1, NUMA
NUA1	Gene	DBP:4590232	0.228	AGTGGAACTCAGTATGAGCTCCG	CAGTTGCTCAGACATATGGGGG	High sulphur keratin, KRN
C-b11098	EST	DBP:458887	0.091	CATTAAGTAGTGGGGGAGAG	CAAGGACAGTGAATAGGG	
RH10753	Gene	DBP:4583588	0.194	GGAGTAGACCATGATTACTG	CATGCTTATTTATCTCG	protein phosphatase 2A, PP2A
EMS1	Gene	DBP:459016	0.84	CGCCTCGGATCTCAGCTACA	GGGCATCAGGGGATGGGTAGA	Amplixin
SHGC-11098	Gene	DBP:4573674	0.137	GCTGCTATCTGCTTGTGATGG	CCGTGGCATAAGATAAGTAAAG	Androgen Receptor
INPL1	Gene	DBP:4572859	0.195	TGGAGTCTGCTGCTGCTGCTG	CAGCACTATCTGCTGGG	51C protein, basophil polyphosphate phosphatase-like 1
D11S2287	EST	DBP:458869	0.1	AAACAGGCTGCTGCTGCTGCTG	GAATGAGGACCACTGGTGAC	NOF1
12461/250	EST	DBP:335210	0.247	TTTCCATATGCTGCTGCTG	CAATGCCAGCTAACAGG	
NDUFV1	EST	DBP:456955	0.19	GAATGCTCTGCTGCTGCTGCTG	CCAGGTGGCTTTCACGG	NDUFV1
AFM032405	MSAT	DBP:609546	0.2	GAACCTCTTTCAGTACGCT	TAATGGTGGCTGCTGCTG	
AFM035949	MSAT	DBP:614025	0.2	GAACCTCTTTCAGTACGCT	TAATGGTGGCTGCTGCTG	
Cda17c12	EST	DBP:458842	0.158	AGGGAATAGTATGCTGCTGCTG	GCAGTGTGTGAAGGCGAG	
SHGC-1384	EST	DBP:4582765	0.137	AGTGGCAAAATGAGAAACAGG	CCACACAGTTCCTCAGTGGC	
RH17410	EST	DBP:4571597	0.128	TGACATCTTTCAGTATGCTG	AGTATCCCACTGATACCG	
RH17414	EST	DBP:4571595	0.121	AGCTGCTGCTGCTGCTGCTG	CAAAAGTGTCTGCTGCTGCTG	
RH17770	EST	DBP:4572301	0.267	SCCTCTCAAGTAGTGTGAGG	IGTGTATCCATAGTGCAGAACAG	
SEA	EST	DBP:4580169	0.13	CTCAGGCCAGGCTACGCT	GGACTCTCCATGCTGCTG	S13 avian erythroblastosis oncogene homolog
RH10889	EST	DBP:4583460	0.107	AAATGATCTCACTGCTGCTG	ACTGAAGAGCTCTGCTGCT	
TIGR-AD08P20	EST	DBP:4587682	0.236	GACATCTGTTAGTCTCATAATG	GGTACAGTGTCTGCTGCT	
TIGR-AD07D16	Gene	DBP:4588398	0.24	CTATGACAAACAGGAGAG	ATCTAGTTCCTCTCTCT	Menin gene (MEN1)
TIGR-AD08B14	EST	DBP:4588882	0.141	GTAATGAGAAACAGCAAAATGA	CTATTGGATGATATGTTATGG	
TIGR-AD08K11	EST	DBP:4589094	0.203	AGTGAACAAATGAGGAGAG	CCTACCCCAAGTACAG	
TIGR-AD08P15	EST	DBP:4589662	0.182	ACTTCTATAAATGAGGTGAG	GAGGACCTCAAGAGGAA	
TIGR-AD08T11	EST	DBP:4589278	0.139	CATACCTCTAGACTCAGGAATC	GAATGATGATCAATCTTCTG	
TIGR-AD08L48	EST	DBP:4589364	0.107	GTGTGAGGAGAAAGCAGT	CTCCAGTAGTCACTTCC	
TIGR-AD08X45	EST	DBP:4589368	0.242	CAAGTACAATAACTTAAGCGG	CAGACCTCTCTCTACAAAAC	
SHGC-11839	Gene	DBP:740339	0.151	TTTATAGAAATGAGCTCTTGCCG	CTACCTCTGCTGCTGCTG	Folate receptor 2 (FRP2)
NIB1242	EST	DBP:3888276	0.149	TCTCATGACAAAGCGTCTG	CCACTGCTCTCTCTCTT	cGMP-stimulated 3'-5'-cyclic nucleotide phosphodiesterase PDE2A3 (PDE2A)
SHGC-13589	Gene	DBP:737568	0.141	CACCAAGGTTGGGGTG	ACTATACGACATGAAGCGGG	Macrophage Migration Inhibitory factor
SHGC-11897	Gene	DBP:874884	0.14	CTAAGTGTGATGACCCG	TGCTCTTCTGAACTTAATCC	P2U Purinceptor
SHGC-15348	EST	DBP:740818	0.141	TCAGGCTCTGCTGCTGAGG	ACATGCTGTGGACCCAT	
Bdab4d05	EST	DBP:445662	0.095	CTGAGCTAGTGCACAG	CCCTGACTGGAGCTGCTC	
Bdab9d07	EST	DBP:445674	0.09	TCAGAGTCTCTGCTGCTG	CAATTCAGCTCATCAGACC	Folate receptor 2 (FRP2)
fol1	Gene	DBP:187640	0.3	CCGCAATTCATCCAGG	GGGTAGGAGGTGCGCAAT	
UBI1738	EST	DBP:828260	0.173	TTCATTTAGGACCTG	CITAAGCCATGTGTTTGG	
W-7351	Gene	DBP:679143	0.324	CTCTCTACACTGCAAAAGC	TGGAGAACCCAGAGGAG	Folate receptor 3 (FRP3)
W-14325	EST	DBP:4576507	0.132	AAAGCAAAAGTAAACAGCAACA	TGGTGTGGGCCCAATATTG	
W-15192	EST	DBP:4576806	0.15	AGAGCACCCTTCTCTCAGCAC	AGAACTCTATCAGCAGGGCG	

TABLE 3: HBM STS Table

WI-17672	EST	GDH:4577492	0.141	AAAAAGGACAGTGTCTAAATTTGA	AATGTTTGTGTTTGTGTTTGTGTT		
SHGC-30732	EST	GDH:4567830	0.105	GAATTAAGGAGTACAGTGCCTGG	TGGTGCCTCAAGAAAG		
SI5G4288	EST	GDH:4566057	0.123	CCATCATCATATTTGTTGACCC	CCAAAGGATGACCAAGTGG		DRES9
WI-13814	EST	GDH:4579280	0.15	TAAGATGCCATTAACCTCATGAC	CTCTGTGCAATGAAGCATCTTACA		Human VEGF related factor isoform VRF188 precursor (VRF)
WI-14122	Gene	GDH:4576114	0.125	CCATCTTTTATCAAGCTGTGG	AGAGGCCATATCATCTCCAC		
27292730	EST	GDH:4598509	0.118	GCAGTGTATTTTCCACAG	GGATGCTTCAGTCCAGAAAG		
SHGC-31328	EST	GDH:4567386	0.15	AGCTTAAAGTAGGACACCATGG	AGATGGCTGCAGGCGCAG		
SGC33858	EST	GDH:4578600	0.127	TGTTGTTATTTCCACCTGGC	TCAGGAAGTAAACACAGGTCTATC		
WI-12161	EST	GDH:1222208	0.15	TTTTTTTTTACAGAAATGAGG	CACAGAGTCCCAGGCTGTGT		
WI-13701	EST	GDH:4574892	0.15	ATGAATCTTACGAAATCCCA	GCCTCAGAGTGTGTGGGT		
WI-14089	EST	GDH:4584373	0.15	MAAGGCCCTTATTAATCTCTCTG	AGCCACAGTCAAGCTTACC		
WI-14272	EST	GDH:4578525	0.125	GGTCTAAGCTTAGAGTCAGCTGG	TGGTCCCACTCACATCCG		
WI-17347	EST	GDH:4578523	0.127	TGGTAAATGATGCCAGA	ATCCGTGTGCTTAGGTGG		
SI5G1561	EST	GDH:4564415	0.215	ACACAGCATCGAGGAGAG	GGAGGCCGCAAGTACTACC		
SI5G1938	EST	GDH:4564568	0.137	GATGGAAGTAGCTCTCTCTGG	GAGGTGCTCTGTGGGGGA		
SI5G2759	EST	GDH:4565137	0.141	CGGTGCTTGGAAAGATG	ACCACCTCACAGCCCTTACA		
RI97	EST	GDH:4559890	0.17	TTACAGCATGAGTCACTACGC	GCCTCAGTAACTTTCAGGGC		
SI5G4784	EST	GDH:4573113	0.141	CCCTCCCTCCAGACACAC	GTGGAATATAGACAGGCGCC		
SI5G4957	EST	GDH:4569051	0.171	AGATAGGGGCAAACTGTG	AGCTGGAACTCTGCTGTCA		
SI5G4974	EST	GDH:4569063	0.166	TTCAGGTTCAGGGCTGTCT	TCCTCTCAGTCTCTCCAGA		
SI5G8144	EST	GDH:4573137	0.17	ACTCAGTCTCTCCACCC	TGGAGGACTGCTTGAGCC		
SI5G9275	EST	GDH:4569999	0.19	GTGATCAGGGCTCAACCTG	TCAAAGTCTGCTGACAGC		Human protein kinase (MLK-3)
SHGC-10687	Gene	GDH:740246	0.277	CTGACGCTGCCCTCACTTTC	CTTAAATGTGTGATGACAAAGC		FGF3
SHGC-11930	Gene	GDH:1231223	0.21	ATTCAGAGCCAGCTGAAA	TACATTTGAACATTTAAACCTGA		FK506-Binding Protein Precursor (FKBP-13)
SHGC-32766	EST	GDH:4567878	0.135	GATCAGCTGTAACGACTGGGA	TGGACAGCTGCTGCTGATGG		4F2 Cell-Surface Antigen Heavy Chain (4F2HC)
FKBP2	EST	GDH:4567878	0.064	ACTGAGCTGTAACTGCTTGG	TGGTCACTGTATTTATGCTAGG		
WI-13116	EST	GDH:4565099	0.202	TATCCCTTTATTTGTTCTCTTGG	CTCATCTCAAGCTGTCTAAGC		
MDU1	Gene	GDH:4565099	0.859	CTTCAAGGCTCTGCTGATAC	TAAAGCCCTATTAACCTCTCC		
S453	Gene	GDH:4565099	0.108	GTGGCTGCAGCTATGTAAAGAC	CAGCAGAGACATGGGCTAAGTCC		
SI51-cSRL-112e11	STS	GDH:195276	0.263	GTCTCCAGAGACAGACAG	GAGAGCAACATTTGCCC		
SI51-cSRL-44e3	STS	GDH:547681	0.135	GATGATTAACCTCTCTGCG	AGAGCTTTGATGACTTGG		Folate receptor2 (FBRP2)
SI51-cSRL-31b12	STS	GDH:456728	0.123	GATGATTAACCTCTCTGCG	GAGCAGCTAAGCACTCATG		Collagen binding protein 2, collagen-2 gene (CBP2)
SI51-cSRL-31b12	STS	GDH:197824	0.166	GTGGTGGTGGGACCTGTGA	TCCCGAGCTCTATCCCACT		Reelin outer segment membrane protein 1, ROM1
SI51-cSRL-31b12	STS	GDH:197824	0.141	GACCAAGTCTGCCAGAG	GTCCAGCTCGTGAATCTCC		
SHGC-10323	Gene	GDH:876135	0.1	GGAGGGATGGCAAGTCTGA	GAAAGGCTTTACATATGG		ZNF128
WI-9219	Gene	GDH:876178	0.172	TCAAAACAGCATCTCTCCA	TCCCTGCTCGGGAAG		
GTC ZNF	Gene	GDH:803197	0.163	CTCAGCAGCCCAAT	GAGAGCAACATTTGCCC		
AFMa152H1	MSAT	GDH:803197	0.263	GTCTCCAGAGACAGACAG	CTCTGTAGGATGAGTGG		
AFMa3312b5	MSAT	GDH:811241	0.151	TATAGACTTCAGGCTGCTGC	GTGAGGCGAGGAATGTGAC		Schellin/serpinase kinase
AFMa3312b5	MSAT	GDH:809521	0.209	TGCTAGCAGCTCTCTACT	CTCCCTGTGCTCCAGTTAT		
AFM212x63	MSAT	GDH:199292	0.13	ATCTAGCAGCAGGAGGCC	AGCAGATCTGCTCTGGAT		
WI-16813	EST		0.262	ACTTTGATTTCCCAAGGA	AAAGATGATGATGGATGGAGC		
WI-19549	EST		0.25	ACAGTGTGATCGGTAGGCA	CCCTATATCTGCTGCTCC		DRES9
WI-20164	EST	GDH:4583084	0.142	GTCCAGTGGGCTTATTT	GAATTCAGGCTCTTGTG		Ultra high-sulphur keratin protein (KRN1)
WI-22393	EST	GDH:4583084	0.214	GTCTAGTGGGAACTCTCAGG	CCAGTACATGGTGGTCAACA		Fas-associated death domain-containing protein, FADD
WI-7687	EST	GDH:1223732	0.273	GGTTTGGTCTGAAGGCAA	GTCTGTGGTGGGGAAG		
EST455579	EST		0.263	GTCTCTTGGAAATTTCTGTT	GGACTGCCCCCTTGAACCTC		
WI-21134	EST		0.25	ATCAAGCTCATCCAGACC	AGACCTGGGGAAGTGGAGAA		
WI-21698	EST	GDH:740192	0.225	ATATAGAGCTGGCAAAATAGG	TAACTTTTGTCAACTCTCTGATT		Atk homeodomain protein, neuroendocrine specific, tk factor
SHGC-7373	STS		0.126	ATTGGCAGCTGGCAAAATAGG	AGGATGTAAAGGTTACAGTGGCG		Chloride channel inducer, ICILN gene
SHGC-36533	STS		0.242	TCAGGCGCTGTGTTGCCGCACTCTG	CCGGAGAGAGACATCAT		
ARX	Gene	GDH:8262613		AGCATGCAAGCTCTTTA	AATGGATGGGGGAATAT		
CLC1PCR	Gene			TCGAGGACAGAAATGCG	AGAGCCGCGTACAGAT		
B1BN21-HL	STS			CTGACGCTATGTGCGC	CCCTCTGTTTCTGTTT		
B23AC17-HR	STS			ATCAGCTGAACTGCCACT	TAGGAGCTTAAGTGGAGAC		
B23G10-HR	STS			CAAGCTTTGAAGGAAG	ACTCCCAAGACTGTGCG		
B37H24-HL	STS			CGCTTCGAGTGGTAA	GACCAGCTGGGAGAAC		
B337L5-HL	STS			CCCTTCTGAGGCAAGAT	GATCAGCTGCAATGAAG		
B362N10-HR	STS			CGCTTCTGAGGCAAGAT	CGCAGGACTGAAGATGA		
B121-HR	STS			CGCTTCTGAGGCAAGAT	TGCTTTCTCTGTGCGGA		
B180D17-HR	STS			TTGAGTACAGCGGCTGAC			
B23656-HR	STS			ACCTGCTCTCTCTCTG			

TABLE 3: HBM STS Table

B278E22-HR	STS	ATGACAGCAAGCATGTT	GTACTGGGATACAGGGC
B312F21-HR	STS	GCAGAGGTCCTTGGAT	TTTGCAGGATCATGCTT
B337H24-HR	STS	CGACATCTTCTCGAGG	ACCTTGCATGTTGTTTT
B358H9-HR	STS	GCATTTCTCTCTCC	TGCTTGTCTCTCTGG
B148N18-HL	STS	ACAGCTCCAGAGAGAGGA	GCAGTCACTGAACACAGA
B172N12-HL	STS	AGGCATCAAGCTTCCCT	GGTTAGAGAACCCGAGCC
B172N12-HL	STS	GTGGTCTGCAAGTACC	GCATCCCTTCTTCCA
B215I11-HR	STS	GACCATTTGACGAGC	GATGGGTGATGAACAA
B223E5-HR	STS	CTCAAGCTTCTGTCATGC	GCTGTGATGCTTGGCT
B312B3-HR	STS	TACAGAAACCGAGCTC	GCCACCAAGGAAGAT
B328G19-HL	STS	AAAGGAGGGATCATGB	TCAGTTAGCAGGAGGAG
B328G19-HR	STS	CTGAGCATCCGATGAGAC	GTGCAAAATGAGCAGCT
B328H10-HL	STS	TCCTAACCTTACTGGGC	TCCTCAACTGGGAATBA
B328H10-HR	STS	TTTACAGAGCACAGGGA	ATCTCCCACTCAAGG
B358G19-HL	STS	GTCCACGGGCTTATCT	TGAGCATAAATTCATAGCTG
B358G19-HR	STS	GGAGAGCAAAATAATCCA	GGTGACAGAAATGTTTCT
B35F16-HL	STS	AGCACCTTATTCATGG	GTAAACACGACGAGACA
B250K11-HR	STS	TCTGCTGCTATATGGAT	GGGGTGAGAGATAGGAA
B338D17-HR	STS	ATGGGATTAATACGGG	AGCTAGCATTGGGCTCT
B268D23-HL	STS	CTGAGGAGAGAGGCTGG	CACAACTCAAGGCAAGTA
B268D23-HR	STS	AGGATGCTCTAAGGTT	CACAACTGCTGGAAGG
B371E15-HR	STS	GGTCTCAGAGGCCCTTA	ACATGCCACCTCTCTACTAA
B372F21-HL	STS	ACTTAACCAAGGATGGG	CAACCCACGAGCATAAGA
B372F21-HL	STS	TAGGGCTCGACTCTGG	ACCCACGAGGCTCTCTC
B338D17-HL	STS	TAAGCGGTGAAGTAG	CTACGGCTCTCTAGGCT
B369H19-HL	STS	TGGGGCCAGATATCTT	CTGGTGTTTGGTGGTGT
B369H19-HR	STS	AGGAAGAGGTCACGAG	CACAACTCCATTTCCA
B444M11-HR	STS	TCATAGGTGATCCACATTT	AAAGTCCACAAAGGCTC
B268L23-HL	STS	GGGTAGGGGATCTTTT	TGTGACATTCATGGC
B250K11-HL	STS	GTCTGGGAAAGATGAA	TCAAAGGCTCTCCATPA
B268L23-HR	STS	TCTTCTGCTACTGGC	TGGGAGTCTCAGAGTATG
B384H4-HL	STS	GGACAGTATGTTGGG	AGGCAGCTGTTTGTGA
B384H4-HR	STS	CTTCTTGAATCCCTGTG	CACCGAATCTCTTAGC
B473O3-HR	STS	GCTGGGAGAGAAATCAAA	GCTTTGCAGAAAGAGCCA
B180D17-HL	STS	ACGCTGTCAGGTACACAT	GGAGGATGCTCAGGTGAT
B200E21-HL	STS	TAGGGGATCTTTTCCA	GAGCAATTTGAAAGCCA
B200E21-HR	STS	ATGGTCAGGCTCTGTGT	ATAGACACGCGCATGCTC
B14L15-HR	STS	ACATTCAGGCTCTGTGT	GCAATCGAAACAGCATTC
B473P8-HR	STS	ATGAGTTGGCAGCTGAAG	AATGAAGGCTTGGCTCC
B189N21-HR	STS	GAAGGAGAGATCCACAGG	TCCTGGGCTACTGAACC
Gene	0.087	CTGAGCTTTGGGCACTGT	CTGCTAGGTGACAGCAGG
B509A5-HL	STS	TGATAGTCTGGAGGCTGT	ACACCTGGCTGAGGAAT
B36F16-HR	STS	GCAGGGGAGTGATATA	TTTGTCTCTTACATGC
B117N18-HL	STS	AAAAATGTGAGCACCTCC	TTTATATTAAGTGCTTGT
B14L16-HL	STS	GTGCAAAAGCCACATAT	AGGAAATGCAAGAGCAG
B21K22-HL	STS	CCACTGAATGCATCTTG	TCGGGTCCAGTCTGCTA
B21K22-HR	STS	AGATTTGGGGAGTCAGG	GCCTCAAGCAATCTC
B223E5-HL	STS	AGCCTCCAGGTGACTAC	GAATCATCAATCCAGCA
B270E22-HL	STS	ATGCTTTACGATTTCC	GAAGGACATGGTCAGCAG
B444M11-HL	STS	GTGCAATTTGTTTCAAA	TGATCCGTGGTAGGGTTA
B543O19-HR	STS	TTTGAAGAAAGAAATGT	TTTATGGGAATTCAGCC
B117N18-HR	STS	CCTTAATGCCCTGATTC	GGCTAGTCTTTCTGAGCC
B543O19-HL	STS	TCAGCTTGTCTTCTCAA	GCCTTACAGCTGAGGA
B442P6-HL	STS	TCAGCTTGTCTTCTCAA	GTAGCCAGCAAGTGTCT
B367H4-HR	STS	CTGGCTGGAGATAGGAT	CTTCCCTCTGGCTATGT
B250E21-HR	STS	GGACCTACTCTTACCA	GGTCTCTTACAGGCA
B250E21-HL	STS	ACCCAGGCTGTGTGT	ACTGAGTAAATTCAGTCCGT
B248C16-HR	STS	GATGCAATTTGCTTACC	TCTGCTTTAGAGCTGTAGC
B248C16-HL	STS	TCAGCTTCAAGAGCAGA	GGAGTACATCCAGGACC
B160D8-HR	STS	TGGTGTCTTTTAAATCCAGA	CTCCCTACTACTTGGCATTG
B539L7-HR	STS	TCTTCCCAAGGGAATCT	TTTATGTCCCTGAGCAC
B473O3-HL	STS	TCCCTGGCTATCTTGAATC	CTTGAAGTGGTCCAGC
AFMa190xd8	D11S4095	GDB:806064	0.193

TABLE 3: HBM STS Table

ARRB1(2)	STS			CGAGAGCCAGTAGATACCA	CATCTCATGCCCTTTCAGT
ARRB1(1)	STS			AGTTCAGAGAGCGAGACGC	CTTGTCTCTCCATGCGTT
P102/3S	STS			GAGCGTAGAGGTTGAGGAG	AAACAAGTCCAGACGCACC
N172A	STS	GDB-6054146	0.208	CTGAACCACTACTGTATGACCTG	CTAACTACTACTCTACAGGCC
N60A	STS	GDB-6054147	0.23	GAAGCATTTCAATACTTTAACTG	CCACTCCAGTGCACCAATC
CH11-44A	STS	GDB-6054148	0.239	CTTCTCTGGCCACTCTGAC	GGTTACCTTTGMAITCCAGC
CN167-2A	STS	GDB-6054149	0.271	TGAGGATGAATGACACATAGG	TTTGGGTCCATTGAGTAGGC
CC11-52AB	STS	GDB-6054150	0.221	AGGGGAAGGAATGTGCTGG	TTCCGCTGAGCGGGCAGTGT
P117F3T	STS	GDB-6054151	0.168	ATTGAAGGTCTCTCAAAAGATGCTG	AGAACGTCAACATACTTTTGGGGGACAC
ARRB1(3)	STS			TTGTATTGAGGAGTTTGTCTG	CGGTACCATCTCTCTCTTC
B215J11-HL	STS		0.122	TTTTTGCTCTCATATGCCC	GGGTGACAGAGCAAGACTCC
B317G1-HR	STS			TTGCTCAAGTTCTCTGG	ACCTGTCTTGAAGGGAG
B317G1-HL	STS			CTTGGCTATTGGACAGC	GGGCATTTACTCAGTTGC
B292J18-HR	STS			CTTGTGTCAGTTGTCAAGG	TGGAATTGTGTGTCTTGG
B10A18-HL	STS			CCAGTCCACGGATGT	ATGGCTGTGTCTCTCA
B10A18-HR	STS			CTGCCCTATCCCTGGACTT	AGTTGCTCTAGTGC
B527D12-HL	STS			CAACAGTCTGCATCCAT	GGATAGTGCACACCCA
B372J11-HR	STS			TGGGTGGTACTATGTTCCAT	AGTTCAGGCCCTTACCAG
B372J11-HL	STS			GGCCACTATCACTCTGTGT	TTTCATGCGGAAGAACAG
B37E17-HR(GS)	STS			ACAGTGACACTAGGAGCGG	TGCCAGGATGAGATAACA
B37E17-HL(GS)	STS			CCGTGGCCACACATACCC	ACAACCAAGATGGAGCCAC
B34F22-HR(GS)	STS			TGCTGTGTAACAAGTCCCA	TGAACGGAGGACCTACCAAG
B34F22-HL(GS)	STS			GCAGGGTCCCACTCACTAAG	GCCTGAGTCCCTTTACGC
B64B22-HR1	STS			ACAGTGGGGACAAGACAGG	TACAGGCGCTCCCTAGTAG
B64B22-HR2	STS			TCCTGTGTTAAGGTTTCCGCC	TGCTCAAACTCCCTCTGC
B64B22-HL	STS			AACATATTCCTCCAGACC	CAGTCCCAAGCAATGAGAAC
B62L11-HL(GS)	STS			CTCTCTGCATGGGAGATC	AGACCTGGGACCACTGTGTG
B62L13-HL(GS)	STS			GGGAGACAGCTCACAGAT	TGATGTTGGGAAGATGTGA
14424-HL	STS			CAGGCATCTTCTATGTCCA	GGGAGGCACAGTCTTCA
B62L11-HR(GS)	STS			ACTTCGTGGCACTGAGTGTG	CCCTTCTTACGGATGAGGA
B65J13-HR(GS)	STS			GCCTGCTGAGCTCTCTGAT	TGGGTCTCTCTGCTGACTT
B62L11-HL2(GS)	STS			TCACCTACTTCCAGCTTCCG	AGACCTGGGACCACTGTGTG
B62L11-HL3(GS)	STS			CTCCTCTGCATGGGAGATC	AATCAGGAGACCTGGGACC

Novel STSs were developed either from publicly available genomic sequence or from sequence-derived BAC insert ends. Primers were chosen using a script which automatically performs vector and repetitive sequence masking using Cross_match (P. Green, U. of Washington) and subsequent primer picking using Primer3 (Rozen, Skaletsky (1996, 1997).

- 5 Primer3 is available at www.genome.wi.mit.edu/genome_software/other/primer3.html.

Polymerase chain reaction (PCR) conditions for each primer pair were initially optimized with respect to $MgCl_2$ concentration. The standard buffer was 10 mM Tris-HCl (pH 8.3), 50 mM KCl, $MgCl_2$, 0.2 mM each dNTP, 0.2 μ M each primer, 2.7 ng/ μ l human DNA, 0.25 U of AmpliTaq (Perkin Elmer) and $MgCl_2$ concentrations of 1.0 mM, 1.5 mM, 10 2.0 mM or 2.4 mM. Cycling conditions included an initial denaturation at 94°C for 2 minutes followed by 40 cycles at 94°C for 15 seconds, 55°C for 25 seconds, and 72°C for 25 seconds followed by a final extension at 72°C for 3 minutes. Depending on the results from the initial round of optimization the conditions were further optimized if necessary. Variables included increasing the annealing temperature to 58°C or 60°C, increasing the 15 cycle number to 42 and the annealing and extension times to 30 seconds, and using AmpliTaqGold (Perkin Elmer).

BAC clones (Kim *et al.*, *Genomics*, 32:213-218 (1996), Shizuya *et al.*, *Proc. Natl. Acad. Sci. USA*, 89:8794-8797 (1992)) containing STS markers of interest were obtained by PCR-based screening of DNA pools from a total human BAC library purchased from 20 Research Genetics. DNA pools derived from library plates 1-596 were used corresponding to nine genomic equivalents of human DNA. The initial screening process involved PCR reactions of individual markers against superpools, i.e., a mixture of DNA derived from all BAC clones from eight 384-well library plates. For each positive superpool, plate (8), row (16) and column (24) pools were screened to identify a unique library address. PCR products

were electrophoresed in 2% agarose gels (Sigma) containing 0.5 µg/ml ethidium bromide in 1X TBE at 150 volts for 45 min. The electrophoresis units used were the Model A3-1 systems from Owl Scientific Products. Typically, gels contained 10 tiers of lanes with 50 wells/tier. Molecular weight markers (100 bp ladder, Life Technologies, Bethesda, MD) were loaded at both ends of the gel. Images of the gels were captured with a Kodak DC40 CCD camera and processed with Kodak 1D software. The gel data were exported as tab delimited text files; names of the files included information about the library screened, the gel image files and the marker screened. These data were automatically imported using a customized Perl script into Filemaker™ PRO (Clariscorp.) databases for data storage and analysis. In cases where incomplete or ambiguous clone address information was obtained, additional experiments were performed to recover a unique, complete library address.

Recovery of clonal BAC cultures from the library involved streaking out a sample from the library well onto LB agar (Maniatis *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1982)) containing 12.5 µg/ml chloramphenicol (Sigma). Two individual colonies and a portion of the initial streak quadrant were tested with appropriate STS markers by colony PCR for verification. Positive clones were stored in LB broth containing 12.5 µg/ml chloramphenicol and 15% glycerol at 70°C.

Several different types of DNA preparation methods were used for isolation of BAC DNA. The manual alkaline lysis miniprep protocol listed below (Maniatis *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1982)) was successfully used for most applications, i.e., restriction mapping, CHEF gel analysis, FISH mapping, but was not successfully reproducible in endsequencing. The

Autogen and Qiagen protocols were used specifically for BAC DNA preparation for endsequencing purposes.

Bacteria were grown in 15 ml Terrific Broth containing 12.5 µg/ml chloramphenicol in a 50 ml conical tube at 37°C for 20 hrs with shaking at 300 rpm. The cultures were

5 centrifuged in a Sorvall RT 6000 D at 3000 rpm (~1800 g) at 4°C for 15 min. The supernatant was then aspirated as completely as possible. In some cases cell pellets were frozen at -20°C at this step for up to 2 weeks. The pellet was then vortexed to homogenize the cells and minimize clumping. 250 µl of P1 solution (50 mM glucose, 15 mM Tris-HCl, pH 8, 10 mM EDTA, and 100 µg/ml RNase A) was added and the mixture pipetted up and

10 down to mix. The mixture was then transferred to a 2 ml Eppendorf tube. 350 µl of P2 solution (0.2 N NaOH, 1% SDS) was then added, the mixture mixed gently and incubated for 5 min. at room temperature. 350 µl of P3 solution (3M KOAc, pH 5.5) was added and the mixture mixed gently until a white precipitate formed. The solution was incubated on ice for 5 min. and then centrifuged at 4°C in a microfuge for 10 min. The supernatant was

15 transferred carefully (avoiding the white precipitate) to a fresh 2 ml Eppendorf tube, and 0.9 ml of isopropanol was added, the solution mixed and left on ice for 5 min. The samples were centrifuged for 10 min., and the supernatant removed carefully. Pellets were washed in 70% ethanol and air dried for 5 min. Pellets were resuspended in 200 µl of TE8 (10 mM Tris-HCl, pH 8.0, 1.0 mM EDTA), and RNase A (Boehringer Mannheim) added to 100 µg/ml.

20 Samples were incubated at 37°C for 30 min., then precipitated by addition of C₂H₃O₂Na·3H₂O to 0.5 M and 2 volumes of ethanol. Samples were centrifuged for 10 min., and the pellets washed with 70% ethanol followed by air drying and dissolving in 50 µl TE8. Typical yields for this DNA prep were 3-5 µg/15 ml bacterial culture. Ten to 15 µl were used

for HindIII restriction analysis; 5 µl was used for NotI digestion and clone insert sizing by CHEF gel electrophoresis.

BACs were inoculated into 15 ml of 2X LB Broth containing 12.5 µg/ml chloramphenicol in a 50 ml conical tube. 4 tubes were inoculated for each clone. Cultures were grown overnight (~16 hr) at 37°C with vigorous shaking (>300 rpm). Standard conditions for BAC DNA isolation were followed as recommended by the Autogen 740 manufacturer. 3 ml samples of culture were placed into Autogen tubes for a total of 60 ml or 20 tubes per clone. Samples were dissolved finally in 100 µl TE8 with 15 seconds of shaking as part of the Autogen protocol. After the Autogen protocol was finished DNA solutions were transferred from each individual tube and pooled into a 2 ml Eppendorf tube. Tubes with large amounts of debris (carry over from the pelleting debris step) were avoided. The tubes were then rinsed with 0.5 ml of TE8 successively and this solution added to the pooled material. DNA solutions were stored at 4°C; clumping tended to occur upon freezing at -20°C. This DNA was either used directly for restriction mapping, CHEF gel analysis or FISH mapping or was further purified as described below for use in endsequencing reactions.

The volume of DNA solutions was adjusted to 2 ml with TE8, samples were then mixed gently and heated at 65°C for 10 min. The DNA solutions were then centrifuged at 4°C for 5 min. and the supernatants transferred to a 15 ml conical tube. The NaCl concentration was then adjusted to 0.75 M (~0.3 ml of 5 M NaCl to the 2 ml sample). The total volume was then adjusted to 6 ml with Qiagen column equilibration buffer (Buffer QBT). The supernatant containing the DNA was then applied to the column and allowed to enter by gravity flow. Columns were washed twice with 10 ml of Qiagen Buffer QC. Bound DNA was then eluted with four separate 1 ml aliquots of Buffer QF kept at 65°C. DNA was precipitated with 0.7 volumes of isopropanol (~2.8 ml). Each sample was then transferred to

4 individual 2.2 ml Eppendorf tubes and incubated at room temperature for 2 hr or overnight. Samples were centrifuged in a microfuge for 10 min. at 4°C. The supernatant was removed carefully and 1 ml of 70% ethanol was added. Samples were centrifuged again and because the DNA pellets were often loose at this stage, the supernatant removed carefully. Samples
5 were centrifuged again to concentrate remaining liquid which was removed with a micropipet tip. DNA pellets were then dried in a desiccator for 10 min. 20 µl of sterile distilled and deionized H₂O was added to each tube which was then placed at 4°C overnight. The four 20 µl samples for each clone were pooled and the tubes rinsed with another 20 µl of sterile distilled and deionized H₂O for a final volume of 100 µl. Samples were then heated at 65°C
10 for 5 min. and then mixed gently. Typical yields were 2-5 µg/60 ml culture as assessed by NotI digestion and comparison with uncut lambda DNA.

3 ml of LB Broth containing 12.5 µg/ml of chloramphenicol was dispensed into autoclaved Autogen tubes. A single tube was used for each clone. For inoculation, glycerol stocks were removed from -70°C storage and placed on dry ice. A small portion of the
15 glycerol stock was removed from the original tube with a sterile toothpick and transferred into the Autogen tube; the toothpick was left in the Autogen tube for at least two minutes before discarding. After inoculation the tubes were covered with tape making sure the seal was tight. When all samples were inoculated, the tube units were transferred into an Autogen rack holder and placed into a rotary shaker at 37°C for 16-17 hours at 250 rpm. Following
20 growth, standard conditions for BAC DNA preparation, as defined by the manufacturer, were used to program the Autogen. Samples were not dissolved in TE8 as part of the program and DNA pellets were left dry. When the program was complete, the tubes were removed from the output tray and 30 µl of sterile distilled and deionized H₂O was added directly to the bottom of the tube. The tubes were then gently shaken for 2-5 seconds and then covered with

parafilm and incubated at room temperature for 1-3 hours. DNA samples were then transferred to an Eppendorf tube and used either directly for sequencing or stored at 4°C for later use.

V. BAC Clone Characterization for Physical Mapping

5 DNA samples prepared either by manual alkaline lysis or the Autogen protocol were digested with HindIII for analysis of restriction fragment sizes. This data were used to compare the extent of overlap among clones. Typically 1-2 µg were used for each reaction. Reaction mixtures included: 1X Buffer 2 (New England Biolabs), 0.1 mg/ml bovine serum albumin (New England Biolabs), 50 µg/ml RNase A (Boehringer Mannheim), and 20 units of
10 HindIII (New England Biolabs) in a final volume of 25 µl. Digestions were incubated at 37°C for 4-6 hours. BAC DNA was also digested with NotI for estimation of insert size by CHEF gel analysis (see below). Reaction conditions were identical to those for HindIII except that 20 units of NotI were used. Six µl of 6X Ficoll loading buffer containing bromphenol blue and xylene cyanol was added prior to electrophoresis.

15 HindIII digests were analyzed on 0.6% agarose (Seakem, FMC Bioproducts) in 1X TBE containing 0.5 µg/ml ethidium bromide. Gels (20 cm X 25 cm) were electrophoresed in a Model A4 electrophoresis unit (Owl Scientific) at 50 volts for 20-24 hrs. Molecular weight size markers included undigested lambda DNA, HindIII digested lambda DNA, and HaeIII digested _X174 DNA. Molecular weight markers were heated at 65°C for 2 min. prior to
20 loading the gel. Images were captured with a Kodak DC40 CCD camera and analyzed with Kodak 1D software.

NotI digests were analyzed on a CHEF DRII (BioRad) electrophoresis unit according to the manufacturer's recommendations. Briefly, 1% agarose gels (BioRad pulsed field

grade) were prepared in 0.5X TBE, equilibrated for 30 minutes in the electrophoresis unit at 14°C, and electrophoresed at 6 volts/cm for 14 hrs with circulation. Switching times were ramped from 10 sec to 20 sec. Gels were stained after electrophoresis in 0.5 µg/ml ethidium bromide. Molecular weight markers included undigested lambda DNA, HindIII digested
5 lambda DNA, lambda ladder PFG ladder, and low range PFG marker (all from New England Biolabs).

BAC DNA prepared either by the manual alkaline lysis or Autogen protocols were labeled for FISH analysis using a Bioprime labeling kit (BioRad) according to the manufacturer's recommendation with minor modifications. Approximately 200 ng of DNA
10 was used for each 50 µl reaction. 3 µl were analyzed on a 2% agarose gel to determine the extent of labeling. Reactions were purified using a Sephadex G50 spin column prior to *in situ* hybridization. Metaphase FISH was performed as described (Ma *et al.*, *Cytogenet. Cell Genet.*, 74:266-271 (1996)).

VI. BAC Endsequencing

15 The sequencing of BAC insert ends utilized DNA prepared by either of the two methods described above. The DYEnamic energy transfer primers and Dynamic Direct cycle sequencing kits from Amersham were used for sequencing reactions. Ready made sequencing mix including the M13 -40 forward sequencing primer was used (Catalog #
US79730) for the T7 BAC vector terminus; ready made sequencing mix (Catalog #
20 US79530) was mixed with the M13 -28 reverse sequencing primer (Catalog # US79339) for the SP6 BAC vector terminus. The sequencing reaction mixes included one of the four fluorescently labeled dye-primers, one of the four dideoxy termination mixes, dNTPs, reaction buffer, and Thermosequenase. For each BAC DNA sample, 3 µl of the BAC DNA

sample was aliquoted to 4 PCR strip tubes. 2 μ l of one of the four dye primer/termination mix combinations was then added to each of the four tubes. The tubes were then sealed and centrifuged briefly prior to PCR. Thermocycling conditions involved a 1 minute denaturation at 95°C, 15 second annealing at 45°C, and extension for 1 minute at 70°C for 35 total cycles.

5 After cycling the plates were centrifuged briefly to collect all the liquid to the bottom of the tubes. 5 μ l of sterile distilled and deionized H₂O was then added into each tube, the plates sealed and centrifuged briefly again. The four samples for each BAC were then pooled together. DNA was then precipitated by adding 1.5 μ l of 7.5 M NH₄OAc and 100 μ l of -20°C 100% ethanol to each tube. Samples were mixed by pipetting up and down once. The

10 plates were then sealed and incubated on ice for 10 minutes. Plates were centrifuged in a table top Haraeus centrifuge at 4000 rpm (3,290 xg) for 30 minutes at 4°C to recover the DNA. The supernatant was removed and excess liquid blotted onto paper towels. Pellets were washed by adding 100 μ l of -20°C 70% ethanol into each tube and recentrifuging at 4000 rpm (3,290 xg) for 10 minutes at 4°C. The supernatant was removed and excess liquid

15 again removed by blotting on a paper towel. Remaining traces of liquid were removed by placing the plates upside down over a paper towel and centrifuging only until the centrifuge reached 800 rpm. Samples were then air dried at room temperature for 30 min. Tubes were capped and stored dry at -20°C until electrophoresis. Immediately prior to electrophoresis the DNA was dissolved in 1.5 μ l of Amersham loading dye. Plates were then sealed and

20 centrifuged at 2000 rpm (825 xg). The plates were then vortexed on a plate shaker for 1-2 minutes. Samples were then recentrifuged at 2000 rpm (825 xg) briefly. Samples were then heated at 65°C for 2 min. and immediately placed on ice. Standard gel electrophoresis was performed on ABI 377 fluorescent sequencers according to the manufacturer's recommendation.

VII. Sub-cloning and Sequencing of HBM BAC DNA

The physical map of the *Zmax1* gene region provides a set of BAC clones that contain within them the *Zmax1* gene and the *HBM* gene. DNA sequencing of several of the BACs from the region has been completed. The DNA sequence data is a unique reagent that includes data that one skilled in the art can use to identify the *Zmax1* gene and the *HBM* gene, or to prepare probes to identify the gene(s), or to identify DNA sequence polymorphisms that identify the gene(s).

BAC DNA was isolated according to one of two protocols, either a Qiagen purification of BAC DNA (Qiagen, Inc. as described in the product literature) or a manual purification which is a modification of the standard alkaline lysis/Cesium Chloride preparation of plasmid DNA (see e.g., Ausubel *et al.*, *Current Protocols in Molecular Biology*, John Wiley & Sons (1997)). Briefly for the manual protocol, cells were pelleted, resuspended in GTE (50 mM glucose, 25 mM Tris-Cl (pH 8), 10 mM EDTA) and lysozyme (50 mg/ml solution), followed by NaOH/SDS (1% SDS/0.2N NaOH) and then an ice-cold solution of 3 M KOAc (pH 4.5-4.8). RnaseA was added to the filtered supernatant, followed by Proteinase K and 20% SDS. The DNA was then precipitated with isopropanol, dried and resuspended in TE (10 mM Tris, 1 mM EDTA (pH 8.0)). The BAC DNA was further purified by Cesium Chloride density gradient centrifugation (Ausubel *et al.*, *Current Protocols in Molecular Biology*, John Wiley & Sons (1997)).

Following isolation, the BAC DNA was sheared hydrodynamically using an HPLC (Hengen, *Trends in Biochem. Sci.*, 22:273-274 (1997)) to an insert size of 2000-3000 bp. After shearing, the DNA was concentrated and separated on a standard 1% agarose gel. A single fraction, corresponding to the approximate size, was excised from the gel and purified

by electroelution (Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring, NY (1989)).

The purified DNA fragments were then blunt-ended using T4 DNA polymerase. The blunt-ended DNA was then ligated to unique BstXI-linker adapters (5'-

5 GTCTTCACCACGGGG and 5' GTGGTGAAGAC in 100-1000 fold molar excess). These linkers were complimentary to the BstXI-cut pMPX vectors (constructed by the inventors), while the overhang was not self-complimentary. Therefore, the linkers would not concatemerize nor would the cut-vector religate itself easily. The linker-adapted inserts were separated from the unincorporated linkers on a 1% agarose gel and purified using GeneClean 10 (BIO 101, Inc.). The linker-adapted insert was then ligated to a modified pBlueScript vector to construct a "shotgun" subclone library. The vector contained an out-of-frame lacZ gene at the cloning site which became in-frame in the event that an adapter-dimer is cloned, allowing these to be avoided by their blue-color.

All subsequent steps were based on sequencing by ABI377 automated DNA 15 sequencing methods. Only major modifications to the protocols are highlighted. Briefly, the library was then transformed into DH5 α competent cells (Life Technologies, Bethesda, MD, DH5 α transformation protocol). It was assessed by plating onto antibiotic plates containing ampicillin and IPTG/Xgal. The plates were incubated overnight at 37°C. Successful transformants were then used for plating of clones and picking for sequencing. The cultures 20 were grown overnight at 37°C. DNA was purified using a silica bead DNA preparation (Ng *et al.*, *Nucl. Acids Res.*, 24:5045-5047 (1996)) method. In this manner, 25 μ g of DNA was obtained per clone.

These purified DNA samples were then sequenced using ABI dye-terminator chemistry. The ABI dye terminator sequence reads were run on ABI377 machines and the

data was directly transferred to UNIX machines following lane tracking of the gels. All reads were assembled using PHRAP (P. Green, Abstracts of DOE Human Genome Program Contractor-Grantee Workshop V, Jan. 1996, p.157) with default parameters and quality scores. The initial assembly was done at 6-fold coverage and yielded an average of 8-15 contigs. Following the initial assembly, missing mates (sequences from clones that only gave one strand reads) were identified and sequenced with ABI technology to allow the identification of additional overlapping contigs. Primers for walking were selected using a Genome Therapeutics program Pick_primer near the ends of the clones to facilitate gap closure. These walks were sequenced using the selected clones and primers. Data were reassembled with PHRAP into sequence contigs.

VIII. Gene Identification by Computational Methods

Following assembly of the BAC sequences into contigs, the contigs were subjected to computational analyses to identify coding regions and regions bearing DNA sequence similarity to known genes. This protocol included the following steps.

1. Degap the contigs: the sequence contigs often contain symbols (denoted by a period symbol) that represent locations where the individual ABI sequence reads have insertions or deletions. Prior to automated computational analysis of the contigs, the periods were removed. The original data was maintained for future reference.
2. BAC vector sequences were "masked" within the sequence by using the program cross match (Phil Green, <http://chimera.biotech.washington.edu/UWGC>). Since the shotgun libraries construction detailed above leaves some BAC vector in the shotgun libraries, this program was used to compare the sequence of the BAC contigs to the BAC

vector and to mask any vector sequence prior to subsequent steps. Masked sequences were marked by an "X" in the sequence files, and remained inert during subsequent analyses.

3. *E. coli* sequences contaminating the BAC sequences were masked by comparing the BAC contigs to the entire *E. coli* DNA sequence.

5 4. Repetitive elements known to be common in the human genome were masked using cross match. In this implementation of crossmatch, the BAC sequence was compared to a database of human repetitive elements (Jerzy Jerka, Genetic Information Research Institute, Palo Alto, CA). The masked repeats were marked by X and remained inert during subsequent analyses.

10 5. The location of exons within the sequence was predicted using the MZEF computer program (Zhang, *Proc. Natl. Acad. Sci.*, 94:565-568 (1997)).

6. The sequence was compared to the publicly available unigene database (National Center for Biotechnology Information, National Library of Medicine, 38A, 8N905, 8600 Rockville Pike, Bethesda, MD 20894; www.ncbi.nlm.nih.gov) using the blastn2
15 algorithm (Altschul *et al.*, *Nucl. Acids Res.*, 25:3389-3402 (1997)). The parameters for this search were: $E=0.05$, $v=50$, $B=50$ (where E is the expected probability score cutoff, V is the number of database entries returned in the reporting of the results, and B is the number of sequence alignments returned in the reporting of the results (Altschul *et al.*, *J. Mol. Biol.*, 215:403-410 (1990))).

20 7. The sequence was translated into protein for all six reading frames, and the protein sequences were compared to a non-redundant protein database compiled from Genpept Swissprot PIR (National Center for Biotechnology Information, National Library of Medicine, 38A, 8N905, 8600 Rockville Pike, Bethesda, MD 20894; www.ncbi.nlm.nih.gov).

The parameters for this search were $E=0.05$, $V=50$, $B=50$, where E , V , and B are defined as above.

8. The BAC DNA sequence was compared to the database of the cDNA clones derived from direct selection experiments (described below) using *blastn2* (Altschul *et al.*,
5 *Nucl. Acids. Res.*, 25:3389-3402 (1997)). The parameters for this search were $E=0.05$,
 $V=250$, $B=250$, where E , V , and B are defined as above.

9. The BAC sequence was compared to the sequences of all other BACs from the HBM region on chromosome 11q12-13 using *blastn2* (Altschul *et al.*, *Nucl. Acids. Res.*,
25:3389-3402 (1997)). The parameters for this search were $E=0.05$, $V=50$, $B=50$, where E , V ,
10 and B are defined as above.

10. The BAC sequence was compared to the sequences derived from the ends of BACs from the HBM region on chromosome 11q12-13 using *blastn2* (Altschul *et al.*, *Nucl. Acids. Res.*, 25:3389-3402 (1997)). The parameters for this search were $E=0.05$, $V=50$, $B=50$, where E , V , and B are defined as above.

11. The BAC sequence was compared to the Genbank database (National Center for Biotechnology Information, National Library of Medicine, 38A, 8N905, 8600 Rockville Pike, Bethesda, MD 20894; www.ncbi.nlm.nih.gov) using *blastn2* (Altschul *et al.*, *Nucl. Acids. Res.*, 25:3389-3402 (1997)). The parameters for this search were $E=0.05$, $V=50$, $B=50$, where E , V , and B are defined as above.

12. The BAC sequence was compared to the STS division of Genbank database (National Center for Biotechnology Information, National Library of Medicine, 38A, 8N905, 8600 Rockville Pike, Bethesda, MD 20894; www.ncbi.nlm.nih.gov) using *blastn2* (Altschul *et al.*, 1997). The parameters for this search were $E=0.05$, $V=50$, $B=50$, where E , V , and B are defined as above.

13. The BAC sequence was compared to the Expressed Sequence (EST) Tag Genbank database (National Center for Biotechnology Information, National Library of Medicine, 38A, 8N905, 8600 Rockville Pike, Bethesda, MD 20894; www.ncbi.nlm.nih.gov) using blastn2 (Altschul *et al.*, *Nucl. Acids. Res.*, 25:3389-3402 (1997)). The parameters for
5 this search were E=0.05, V=250, B=250, where E, V, and B are defined as above.

IX. Gene Identification by Direct cDNA Selection

Primary linkered cDNA pools were prepared from bone marrow, calvarial bone, femoral bone, kidney, skeletal muscle, testis and total brain. Poly (A) + RNA was prepared from calvarial and femoral bone tissue (Chomczynski *et al.*, *Anal. Biochem.*, 162:156-159
10 (1987); D'Alessio *et al.*, *Focus*, 9:1-4 (1987)) and the remainder of the mRNA was purchased from Clontech (Palo Alto, California). In order to generate oligo(dT) and random primed cDNA pools from the same tissue, 2.5 µg mRNA was mixed with oligo(dT) primer in one reaction and 2.5 µg mRNA was mixed with random hexamers in another reaction, and both were converted to first and second strand cDNA according to manufacturers
15 recommendations (Life Technologies, Bethesda, MD). Paired phosphorylated cDNA linkers (see sequence below) were annealed together by mixing in a 1:1 ratio (10 µg each) incubated at 65°C for five minutes and allowed to cool to room temperature.

Paired linkers oligo1/2

OLIGO 1: 5'CTG AGC GGA ATT CGT GAG ACC3' (SEQ ID NO:12)

20 OLIGO 2: 5'TTG GTC TCA CGT ATT CCG CTC GA3' (SEQ ID NO:13)

Paired linkers oligo3/4

OLIGO 3: 5'CTC GAG AAT TCT GGA TCC TC3' (SEQ ID NO:14)

OLIGO 4: 5'TTG AGG ATC CAG AAT TCT CGA G3' (SEQ ID NO:15)

Paired linkers oligo5/6

OLIGO 5: 5'TGT ATG CGA ATT CGC TGC GCG3' (SEQ ID NO:16)

OLIGO 6: 5'TTC GCG CAG CGA ATT CGC ATA CA3' (SEQ ID NO:17)

5 Paired linkers oligo7/8

OLIGO 7: 5'GTC CAC TGA ATT CTC AGT GAG3' (SEQ ID NO:18)

OLIGO 8: 5'TTG TCA CTG AGA ATT CAG TGG AC3' (SEQ ID NO:19)

Paired linkers oligo11/12

OLIGO 11: 5'GAA TCC GAA TTC CTG GTC AGC3' (SEQ ID NO:20)

10 OLIGO 12: 5'TTG CTG ACC AGG AAT TCG GAT TC3' (SEQ ID NO:21)

Linkers were ligated to all oligo(dT) and random primed cDNA pools (see below) according to manufacturers instructions (Life Technologies, Bethesda, MD).

Oligo 1/2 was ligated to oligo(dT) and random primed cDNA pools prepared from bone marrow. Oligo 3/4 was ligated to oligo(dT) and random primed cDNA pools prepared from calvarial bone. Oligo 5/6 was ligated to oligo(dT) and random primed cDNA pools prepared from brain and skeletal muscle. Oligo 7/8 was ligated to oligo(dT) and random primed cDNA pools prepared from kidney. Oligo 11/12 was ligated to oligo(dT) and random primed cDNA pools prepared from femoral bone.

The cDNA pools were evaluated for length distribution by PCR amplification using 1
20 µl of a 1:1, 1:10, and 1:100 dilution of the ligation reaction, respectively. PCR reactions were performed in a Perkin Elmer 9600, each 25 µl volume reaction contained 1 µl of DNA, 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM MgCl₂, 0.001% gelatin, 200 mM each

dNTPs, 10 μ M primer and 1 unit Taq DNA polymerase (Perkin Elmer) and was amplified under the following conditions: 30 seconds at 94°C, 30 seconds at 60°C and 2 minutes at 72°C for 30 cycles. The length distribution of the amplified cDNA pools were evaluated by electrophoresis on a 1% agarose gel. The PCR reaction that gave the best representation of the random primed and oligo(dT) primed cDNA pools was scaled up so that ~2-3 μ g of each cDNA pool was produced. The starting cDNA for the direct selection reaction comprised of 0.5 μ g of random primed cDNAs mixed with 0.5 μ g of oligo(dT) primed cDNAs.

The DNA from the 54 BACs that were used in the direct cDNA selection procedure was isolated using Nucleobond AX columns as described by the manufacturer (The Nest Group, Inc.).

The BACs were pooled in equimolar amounts and 1 μ g of the isolated genomic DNA was labelled with biotin 16-UTP by nick translation in accordance with the manufacturers instructions (Boehringer Mannheim). The incorporation of the biotin was monitored by methods that could be practiced by one skilled in the art (Del Mastro and Lovett, *Methods in Molecular Biology*, Humana Press Inc., NJ (1996)).

Direct cDNA selection was performed using methods that could be practiced by one skilled in the art (Del Mastro and Lovett, *Methods in Molecular Biology*, Humana Press Inc., NJ (1996)). Briefly, the cDNA pools were multiplexed in two separate reactions: In one reaction cDNA pools from bone marrow, calvarial bone, brain and testis were mixed, and in the other cDNA pools from skeletal muscle, kidney and femoral bone were mixed.

Suppression of the repeats, yeast sequences and plasmid in the cDNA pools was performed to a Cot of 20. 100 ng of biotinylated BAC DNA was mixed with the suppressed cDNAs and hybridized in solution to a Cot of 200. The biotinylated DNA and the cognate cDNAs was captured on streptavidin-coated paramagnetic beads. The beads were washed and the primary

selected cDNAs were eluted. These cDNAs were PCR amplified and a second round of direct selection was performed. The product of the second round of direct selection is referred to as the secondary selected material. A Galanin cDNA clone, previously shown to map to 11q12-13 (Evans, *Genomics*, 18:473-477 (1993)), was used to monitor enrichment

5 during the two rounds of selection.

The secondary selected material from bone marrow, calvarial bone, femoral bone, kidney, skeletal muscle, testis and total brain was PCR amplified using modified primers of oligos 1, 3, 5, 7 and 11, shown below, and cloned into the UDG vector pAMP10 (Life Technologies, Bethesda, MD), in accordance with the manufacturer's recommendations.

10 Modified primer sequences:

Oligo1-CUA: 5'CUA CUA CUA CUA CTG AGC GGA ATT CGT GAG ACC3' (SEQ ID NO:22)

Oligo3-CUA: 5'CUA CUA CUA CUA CTC GAG AAT TCT GGA TCC TC3' (SEQ ID NO:23)

15 Oligo5-CUA: 5'CUA CUA CUA CUA TGT ATG CGA ATT CGC TGC GCG3' (SEQ ID NO:24)

Oligo7-CUA: 5'CUA CUA CUA CUA GTC CAC TGA ATT CTC AGT GAG3' (SEQ ID NO:25)

20 Oligo11-CUA: 5'CUA CUA CUA CUA GAA TCC GAA TTC CTG GTC AGC3' (SEQ ID NO:26)

The cloned secondary selected material, from each tissue source, was transformed into MAX Efficiency DH5a Competent Cells (Life Technologies, Bethesda, MD) as recommended by the manufacturer. 384 colonies were picked from each transformed source and arrayed into four 96 well microtiter plates.

All secondarily selected cDNA clones were sequenced using M13 dye primer terminator cycle sequencing kit (Applied Biosystems), and the data collected by the ABI 377 automated fluorescence sequencer (Applied Biosystems).

All sequences were analyzed using the BLASTN, BLASTX and FASTA programs (Altschul *et al.*, *J. Mol. Biol.*, 215:403-410 (1990), Altschul *et al.*, *Nucl. Acids. Res.*, 25:3389-3402 (1997)). The cDNA sequences were compared to a database containing sequences derived from human repeats, mitochondrial DNA, ribosomal RNA, *E. coli* DNA to remove background clones from the dataset using the program cross_match. A further round of comparison was also performed using the program BLASTN2 against known genes (Genbank) and the BAC sequences from the HBM region. Those cDNAs that were >90% homologous to these sequences were filed according to the result and the data stored in a database for further analysis. cDNA sequences that were identified but did not have significant similarity to the BAC sequences from the HBM region or were eliminated by cross_match were hybridized to nylon membranes which contained the BACs from the HBM region, to ascertain whether they hybridized to the target.

Hybridization analysis was used to map the cDNA clones to the BAC target that selected them. The BACs that were identified from the HBM region were arrayed and grown into a 96 well microtiter plate. LB agar containing 25 µg/ml kanamycin was poured into 96 well microtiter plate lids. Once the agar had solidified, pre-cut Hybond N+ nylon membranes (Amersham) were laid on top of the agar and the BACs were stamped onto the membranes in duplicate using a hand held 96 well replica plater (V&P Scientific, Inc.). The plates were incubated overnight at 37°C. The membranes were processed according to the manufacturers recommendations.

The cDNAs that needed to be mapped by hybridization were PCR amplified using the relevant primer (oligos 1, 3, 5, 7 and 11) that would amplify that clone. For this PCR amplification, the primers were modified to contain a linkered digoxigenin molecule at the 5' of the oligonucleotide. The PCR amplification was performed under the same conditions as described in Preparation of cDNA Pools (above). The PCR products were evaluated for quality and quantity by electrophoresis on a 1% agarose gel by loading 5 μ l of the PCR reaction. The nylon membranes containing the stamped BACs were individually pre-hybridized in 50 ml conical tubes containing 10 ml of hybridization solution (5x SSPE, 0.5x Blotto, 2.5% SDS and 1 mM EDTA (pH 8.0)). The 50 ml conical tubes were placed in a rotisserie oven (Robbins Scientific) for 2 hours at 65°C. 25 ng of each cDNA probe was denatured and added into individual 50 ml conical tubes containing the nylon membrane and hybridization solution. The hybridization was performed overnight at 65°C. The filters were washed for 20 minutes at 65°C in each of the following solutions: 3x SSPE, 0.1% SDS; 1x SSPE, 0.1% SDS and 0.1x SSPE, 0.1% SDS.

The membranes were removed from the 50 ml conical tubes and placed in a dish. Acetate sheets were placed between each membrane to prevent them from sticking to each other. The incubation of the membranes with the Anti-DIG-AP and CDP-Star was performed according to manufacturers recommendations (Boehringer Mannheim). The membranes were wrapped in Saran wrap and exposed to Kodak Bio-Max X-ray film for 1 hour.

20 X. cDNA Cloning and Expression Analysis

To characterize the expression of the genes identified by direct cDNA selection and genomic DNA sequencing in comparison to the publicly available databases, a series of experiments were performed to further characterize the genes in the HBM region. First,

oligonucleotide primers were designed for use in the polymerase chain reaction (PCR) so that portions of a cDNA, EST, or genomic DNA could be amplified from a pool of DNA molecules (a cDNA library) or RNA population (RT-PCR and RACE). The PCR primers were used in a reaction containing genomic DNA to verify that they generated a product of the size predicted based on the genomic (BAC) sequence. A number of cDNA libraries were then examined for the presence of the specific cDNA or EST. The presence of a fragment of a transcription unit in a particular cDNA library indicates a high probability that additional portions of the same transcription unit will be present as well.

A critical piece of data that is required when characterizing novel genes is the length, in nucleotides, of the processed transcript or messenger RNA (mRNA). One skilled in the art primarily determines the length of an mRNA by Northern blot hybridization (Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor NY (1989)). Groups of ESTs and direct-selected cDNA clones that displayed significant sequence similarity to sequenced BACs in the critical region were grouped for convenience into approximately 30 kilobase units. Within each 30 kilobase unit there were from one up to fifty ESTs and direct-selected cDNA clones which comprised one or more independent transcription units. One or more ESTs or direct-selected cDNAs were used as hybridization probes to determine the length of the mRNA in a variety of tissues, using commercially available reagents (Multiple Tissue Northern blot; Clontech, Palo Alto, California) under conditions recommended by the manufacturer.

Directionally cloned cDNA libraries from femoral bone, and calvarial bone tissue were constructed by methods familiar to one skilled in the art (for example, Soares in *Automated DNA Sequencing and Analysis*, Adams, Fields and Venter, Eds., Academic Press, NY, pages 110-114 (1994)). Bones were initially broken into fragments with a hammer, and

the small pieces were frozen in liquid nitrogen and reduced to a powder in a tissue pulverizer (Spectrum Laboratory Products). RNA was extracted from the powdered bone by homogenizing the powdered bone with a standard Acid Guanidinium Thiocyanate-Phenol-Chloroform extraction buffer (e.g. Chomczynski and Sacchi, *Anal. Biochem.*, 162:156-159 (1987)) using a polytron homogenizer (Brinkman Instruments). Additionally, human brain and lung total RNA was purchased from Clontech. PolyA RNA was isolated from total RNA using dynabeads-dT according to the manufacturer's recommendations (Dynal, Inc.).

First strand cDNA synthesis was initiated using an oligonucleotide primer with the sequence: 5'-AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTTT-3' (SEQ ID NO:27). This primer introduces a NotI restriction site (underlined) at the 3' end of the cDNA. First and second strand synthesis were performed using the "one-tube" cDNA synthesis kit as described by the manufacturer (Life Technologies, Bethesda, MD). Double stranded cDNAs were treated with T4 polynucleotide kinase to ensure that the ends of the molecules were blunt (Soares, in *Automated DNA Sequencing and Analysis*, Adams, Fields and Venter, Eds., Academic Press, NY, pages 110-114 (1994)), and the blunt ended cDNAs were then size selected by a Biogel column (Huynh *et al.* in *DNA Cloning*, Vol. 1, Glover, Ed., IRL Press, Oxford, pages 49-78 (1985)) or with a size-sep 400 sepharose column (Pharmacia, catalog # 27-5105-01). Only cDNAs of 400 base pairs or longer were used in subsequent steps. EcoRI adapters (sequence: 5' OH-AATTCGGCACGAG-OH 3' (SEQ ID NO:28), and 5' p-CTCGTGCCG-OH 3' (SEQ ID NO:29)) were then ligated to the double stranded cDNAs by methods familiar to one skilled in the art (Soares, 1994). The EcoRI adapters were then removed from the 3' end of the cDNA by digestion with NotI (Soares, 1994). The cDNA was then ligated into the plasmid vector pBluescript II KS+ (Stratagene,

La Jolla, California), and the ligated material was transformed into *E. coli* host DH10B or DH12S by electroporation methods familiar to one skilled in the art (Soares, 1994). After growth overnight at 37°C, DNA was recovered from the *E. coli* colonies after scraping the plates by processing as directed for the Mega-prep kit (Qiagen, Chatsworth, California). The
5 quality of the cDNA libraries was estimated by counting a portion of the total numbers of primary transformants and determining the average insert size and the percentage of plasmids with no cDNA insert. Additional cDNA libraries (human total brain, heart, kidney, leukocyte, and fetal brain) were purchased from Life Technologies, Bethesda, MD.

cDNA libraries, both oligo (dT) and random hexamer (N_6) primed, were used for
10 isolating cDNA clones transcribed within the HBM region: human bone, human brain, human kidney and human skeletal muscle (all cDNA libraries were made by the inventors, except for skeletal muscle (dT) and kidney (dT) cDNA libraries). Four 10 x 10 arrays of each of the cDNA libraries were prepared as follows: the cDNA libraries were titrated to 2.5×10^6 using primary transformants. The appropriate volume of frozen stock was used to inoculate 2 L of
15 LB/ampicillin (100 mg/ml). This inoculated liquid culture was aliquotted into 400 tubes of 4 ml each. Each tube contained approximately 5000 cfu. The tubes were incubated at 30°C overnight with gentle agitation. The cultures were grown to an OD of 0.7-0.9. Frozen stocks were prepared for each of the cultures by aliquotting 100 μ l of culture and 300 μ l of 80% glycerol. Stocks were frozen in a dry ice/ethanol bath and stored at -70°C. The remaining
20 culture was DNA prepared using the Qiagen (Chatsworth, CA) spin miniprep kit according to the manufacturer's instructions. The DNAs from the 400 cultures were pooled to make 80 column and row pools. The cDNA libraries were determined to contain HBM cDNA clones of interest by PCR. Markers were designed to amplify putative exons. Once a standard PCR optimization was performed and specific cDNA libraries were determined to contain cDNA

clones of interest, the markers were used to screen the arrayed library. Positive addresses indicating the presence of cDNA clones were confirmed by a second PCR using the same markers.

Once a cDNA library was identified as likely to contain cDNA clones corresponding
5 to a specific transcript of interest from the HBM region, it was manipulated to isolate the clone or clones containing cDNA inserts identical to the EST or direct-selected cDNA of interest. This was accomplished by a modification of the standard "colony screening" method (Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor NY (1989)). Specifically, twenty 150 mm LB+ampicillin
10 agar plates were spread with 20,000 colony forming units (cfu) of cDNA library and the colonies allowed to grow overnight at 37°C. Colonies were transferred to nylon filters (Hybond from Amersham, or equivalent) and duplicates prepared by pressing two filters together essentially as described (Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor NY (1989)). The "master" plate was
15 then incubated an additional 6-8 hours to allow the colonies to grow back. The DNA from the bacterial colonies was then affixed to the nylon filters by treating the filters sequentially with denaturing solution (0.5 N NaOH, 1.5 M NaCl) for two minutes, neutralization solution (0.5 M Tris-Cl pH 8.0, 1.5 M NaCl) for two minutes (twice). The bacterial colonies were removed from the filters by washing in a solution of 2X SSC/0.1% SDS for one minute while
20 rubbing with tissue paper. The filters were air dried and baked under vacuum at 80°C for 1-2 hours.

A cDNA hybridization probe was prepared by random hexamer labeling (Fineberg and Vogelstein, *Anal. Biochem.*, 132:6-13 (1983)) or by including gene-specific primers and no random hexamers in the reaction (for small fragments). Specific activity was calculated

and was $>5 \times 10^8$ cpm/ 10^8 μ g of cDNA. The colony membranes were then prewashed in 10 mM Tris-Cl pH 8.0, 1 M NaCl, 1 mM EDTA, 0.1% SDS for 30 minutes at 55 °C. Following the prewash, the filters were prehybridized in > 2 ml/filter of 6X SSC, 50 % deionized formamide, 2% SDS, 5X Denhardt's solution, and 100 mg/ml denatured salmon sperm DNA, at 42 °C for 30 minutes. The filters were then transferred to hybridization solution (6X SSC, 2% SDS, 5X Denhardt's, 100 mg/ml denatured salmon sperm DNA) containing denatured α^{32} P-dCTP-labelled cDNA probe and incubated at 42 °C for 16-18 hours.

After the 16-18 hour incubation, the filters were washed under constant agitation in 2X SSC, 2% SDS at room temperature for 20 minutes, followed by two washes at 65 °C for 15 minutes each. A second wash was performed in 0.5 X SSC, 0.5% SDS for 15 minutes at 65 °C. Filters were then wrapped in plastic wrap and exposed to radiographic film for several hours to overnight. After film development, individual colonies on plates were aligned with the autoradiograph so that they could be picked into a 1 ml solution of LB Broth containing ampicillin. After shaking at 37 °C for 1-2 hours, aliquots of the solution were plated on 150 mm plates for secondary screening. Secondary screening was identical to primary screening (above) except that it was performed on plates containing ~250 colonies so that individual colonies could be clearly identified for picking.

After colony screening with radiolabeled probes yielded cDNA clones, the clones were characterized by restriction endonuclease cleavage, PCR, and direct sequencing to confirm the sequence identity between the original probe and the isolated clone. To obtain the full-length cDNA, the novel sequence from the end of the clone identified was used to probe the library again. This process was repeated until the length of the cDNA cloned matches that estimated to be full-length by the northern blot analysis.

RT-PCR was used as another method to isolate full length clones. The cDNA was synthesized and amplified using a "Superscript One Step RT-PCR" kit (Life Technologies, Gaithersburg, MD). The procedure involved adding 1.5 µg of RNA to the following: 25 µl of reaction mix provided which is a proprietary buffer mix with MgSO₄ and dNTP's, 1 µl sense primer (10 µM) and 1 µl anti-sense primer (10 µM), 1 µl reverse transcriptase and Taq DNA polymerase mix provided and autoclaved water to a total reaction mix of 50 µl. The reaction was then placed in a thermocycler for 1 cycle at 50°C for 15 to 30 minutes, then 94°C for 15 seconds, 55-60°C for 30 seconds and 68-72°C for 1 minute per kilobase of anticipated product and finally 1 cycle of 72°C for 5-10 minutes. The sample was analyzed on an agarose gel. The product was excised from the gel and purified from the gel (GeneClean, Bio 101). The purified product was cloned in pCTNR (General Contractor DNA Cloning System, 5 Prime - 3 Prime, Inc.) and sequenced to verify that the clone was specific to the gene of interest.

Rapid Amplification of cDNA ends (RACE) was performed following the manufacturer's instructions using a Marathon cDNA Amplification Kit (Clontech, Palo Alto, CA) as a method for cloning the 5' and 3' ends of candidate genes. cDNA pools were prepared from total RNA by performing first strand synthesis, where a sample of total RNA sample was mixed with a modified oligo (dT) primer, heated to 70°C, cooled on ice and followed by the addition of: 5x first strand buffer, 10 mM dNTP mix, and AMV Reverse Transcriptase (20 U/µl). The tube was incubated at 42°C for one hour and then the reaction tube was placed on ice. For second strand synthesis, the following components were added directly to the reaction tube: 5x second strand buffer, 10 mM dNTP mix, sterile water, 20x second strand enzyme cocktail and the reaction tube was incubated at 16°C for 1.5 hours. T4 DNA Polymerase was added to the reaction tube and incubated at 16°C for 45 minutes. The

second-strand synthesis was terminated with the addition of an EDTA/Glycogen mix. The sample was subjected to a phenol/chloroform extraction and an ammonium acetate precipitation. The cDNA pools were checked for quality by analyzing on an agarose gel for size distribution. Marathon cDNA adapters (Clontech) were then ligated onto the cDNA ends. The specific adapters contained priming sites that allowed for amplification of either 5' or 3' ends, depending on the orientation of the gene specific primer (GSP) that was chosen. An aliquot of the double stranded cDNA was added to the following reagents: 10 μ M Marathon cDNA adapter, 5x DNA ligation buffer, T4 DNA ligase. The reaction was incubated at 16°C overnight. The reaction was heat inactivated to terminate the reaction.

10 PCR was performed by the addition of the following to the diluted double stranded cDNA pool: 10x cDNA PCR reaction buffer, 10 μ M dNTP mix, 10 μ M GSP, 10 μ M AP1 primer (kit), 50x Advantage cDNA Polymerase Mix. Thermal Cycling conditions were 94°C for 30 seconds, 5 cycles of 94°C for 5 seconds, 72°C for 4 minutes, 5 cycles of 94°C for 5 seconds, 70°C for 4 minutes, 23 cycles of 94°C for 5 seconds, 68°C for 4 minutes. After the first

15 round of PCR was performed using the GSP to extend to the end of the adapter to create the adapter primer binding site, exponential amplification of the specific cDNA of interest was observed. Usually a second nested PCR is performed to confirm the specific cDNA. The RACE product was analyzed on an agarose gel and then excised and purified from the gel (GeneClean, BIO 101). The RACE product was then cloned into pCTNR (General

20 Contractor DNA Cloning System, 5' - 3', Inc.) and the DNA sequence determined to verify that the clone is specific to the gene of interest.

XI. Mutation Analysis

Comparative genes were identified using the above procedures and the exons from each gene were subjected to mutation detection analysis. Comparative DNA sequencing was used to identify polymorphisms in HBM candidate genes from chromosome 11q12-13. DNA
5 sequences for candidate genes were amplified from patient lymphoblastoid cell lines.

The inventors developed a method based on analysis of direct DNA sequencing of PCR products amplified from candidate regions to search for the causative polymorphism. The procedure consisted of three stages that used different subsets of HBM family to find segregating polymorphisms and a population panel to assess the frequency of the
10 polymorphisms. The family resources result from a single founder leading to the assumption that all affected individuals will share the same causative polymorphism.

Candidate regions were first screened in a subset of the HBM family consisting of the proband, daughter, and her mother, father and brother. Monochromosomal reference sequences were produced concurrently and used for comparison. The mother and daughter
15 carried the HBM polymorphism in this nuclear family, providing the ability to monitor polymorphism transmission. The net result is that two HBM chromosomes and six non-HBM chromosomes were screened. This allowed exclusion of numerous frequent alleles. Only alleles exclusively present in the affected individuals passed to the next level of analysis.

20 Polymorphisms that segregated exclusively with the HBM phenotype in this original family were then re-examined in an extended portion of the HBM pedigree consisting of two additional nuclear families. These families consisted of five HBM and three unaffected individuals. The HBM individuals in this group included the two critical crossover individuals, providing the centromeric and telomeric boundaries of the critical region.

Tracking the heredity of polymorphisms between these individuals and their affected parents allowed for further refining of the critical region. This group brought the total of HBM chromosomes screened to seven and the total of non-HBM chromosomes to seventeen.

When a given polymorphism continued to segregate exclusively with the HBM phenotype in the extended group, a population panel was then examined. This panel of 84 persons consisted of 42 individuals known to have normal bone mineral density and 42 individuals known to be unrelated but with untyped bone mineral density. Normal bone mineral density is within two standard deviations of BMD Z score 0. The second group was from the widely used CEPH panel of individuals. Any segregating polymorphisms found to be rare in this population were subsequently examined on the entire HBM pedigree and a larger population.

Polymerase chain reaction (PCR) was used to generate sequencing templates from the HBM family's DNA and monochromosomal controls. Enzymatic amplification of genes within the HBM region on 11q12-13 was accomplished using the PCR with oligonucleotides flanking each exon as well as the putative 5' regulatory elements of each gene. The primers were chosen to amplify each exon as well as 15 or more base pairs within each intron on either side of the splice. All PCR primers were made as chimeras to facilitate dye primer sequencing. The M13-21F (5'- GTA A CGA CGG CCA GT -3') (SEQ ID NO:30) and -28REV (5'- AAC AGC TAT GAC CAT G -3') (SEQ ID NO:31) primer binding sites were built on to the 5' end of each forward and reverse PCR primer, respectively, during synthesis. 150 ng of genomic DNA was used in a 50 µl PCR with 2U AmpliTaq, 500 nM primer and 125 µM dNTP. Buffer and cycling conditions were specific to each primer set. TaqStart antibody (Clontech) was used for hot start PCR to minimize primer dimer formation. 10% of

the product was examined on an agarose gel. The appropriate samples were diluted 1:25 with deionized water before sequencing.

Each PCR product was sequenced according to the standard Energy Transfer primer (Amersham) protocol. All reactions took place in 96 well trays. 4 separate reactions, one
5 each for A, C, G and T were performed for each template. Each reaction included 2 μ l of the sequencing reaction mix and 3 μ l of diluted template. The plates were then heat sealed with foil tape and placed in a thermal cycler and cycled according to the manufacturer's recommendation. After cycling, the 4 reactions were pooled. 3 μ l of the pooled product was transferred to a new 96 well plate and 1 μ l of the manufacturer's loading dye was added to
10 each well. All 96 well pipetting procedures occurred on a Hydra 96 pipetting station (Robbins Scientific, USA). 1 μ l of pooled material was directly loaded onto a 48 lane gel running on an ABI 377 DNA sequencer for a 10 hour, 2.4 kV run.

Polyphred (University of Washington) was used to assemble sequence sets for viewing with *Consed* (University of Washington). Sequences were assembled in groups
15 representing all relevant family members and controls for a specified target region. This was done separately for each of the three stages. Forward and reverse reads were included for each individual along with reads from the monochromosomal templates and a color annotated reference sequence. *Polyphred* indicated potential polymorphic sites with a purple flag. Two readers independently viewed each assembly and assessed the validity of the purple-flagged
20 sites.

A total of 23 exons present in the mature mRNA and several other portions of the primary transcript were evaluated for heterozygosity in the nuclear family of two HBM-affected and two unaffected individuals. Twenty-five single nucleotide polymorphisms (SNPs) were identified, as shown in the table below.

TABLE 4: Single Nucleotide Polymorphisms in the *Zmax1* gene and Environs

	Exon Name	Location	Base Change
	b200e21-h_Contig1_1.nt	69169 (309G)	C/A
	b200e21-h_Contig4_12.nt	27402 (309G)	A/G
5	b200e21-h_Contig4_13.nt	27841 (309G)	T/C
	b200e21-h_Contig4_16.nt	35600 (309G)	A/G
	b200e21-h_Contig4_21.nt	45619 (309G)	G/A
	b200e21-h_Contig4_22.nt-a	46018 (309G)	T/G
	b200e21-h_Contig4_22.nt-b	46093 (309G)	T/G
10	b200e21-h_Contig4_22.nt-c	46190 (309G)	A/G
	b200e21-h_Contig4_24.nt-a	50993 (309G)	T/C
	b200e21-h_Contig4_24.nt-b	51124 (309G)	C/T
	b200e21-h_Contig4_25.nt	55461 (309G)	C/T
	b200e21-h_Contig4_33.nt-a	63645 (309G)	C/A
15	b200e21-h_Contig4_33.nt-b	63646 (309G)	A/C
	b200e21-h_Contig4_61.nt	24809 (309G)	T/G
	b200e21-h_Contig4_62.nt	27837 (309G)	T/C
	b200e21-h_Contig4_63.nt-a	31485 (309G)	C/T
	b200e21-h_Contig4_63.nt-b	31683 (309G)	A/G
20	b200e21-h_Contig4_9.nt	24808 (309G)	T/G
	b527d12-h_Contig030g_1.nt-a	31340 (308G)	T/C
	b527d12-h_Contig030g_1.nt-b	32538 (308G)	A/G
	b527d12-h_Contig080C_2.nt	13224 (308G)	A/G
	b527d12-h_Contig087C_1.nt	21119 (308G)	C/A
25	b527d12-h_Contig087C_4.nt	30497 (308G)	G/A
	b527d12-h_Contig088C_4.nt	24811 (309G)	A/C
	b527d12-h_Contig089_1HP.nt	68280 (309G)	G/A

In addition to the polymorphisms presented in Table 4, two additional polymorphisms can also be present in SEQ ID NO:2. These is a change at position 2002 of SEQ ID NO:2. Either a guanine or an adenine can appear at this position. This polymorphism is silent and is not associated with any change in the amino acid sequence.

- 5 The second change is at position 4059 of SEQ ID NO:2 corresponding in a cytosine (C) to thymine (T) change. This polymorphism results in a corresponding amino acid change from a valine (V) to an alanine (A). Other polymorphisms were found in the candidate gene exons and adjacent intron sequences. Any one or combination of the polymorphisms listed in Table 4 or the two discussed above could also have a minor effect on bone mass or lipid levels
- 10 when present in SEQ ID NO:2.

The present invention encompasses the nucleic acid sequences having the nucleic acid sequence of SEQ ID NO: 1 with the above-identified point mutations.

Preferably, the present invention encompasses the nucleic acid of SEQ ID NO: 2.

- Specifically, a base-pair substitution changing G to T at position 582 in the coding sequence of Zmax1 (the *HBM* gene) was identified as heterozygous in all HBM individuals, and not
- 15 found in the unaffected individuals (i.e., b527d12-h_Contig087C_1.nt). Fig. 5 shows the order of the contigs in B527D12. The direction of transcription for the *HBM* gene is from left to right. The sequence of contig308G of B527D12 is the reverse complement of the coding region to the *HBM* gene. Therefore, the relative polymorphism in contig 308G shown in
- 20 Table 4 as a base change substitution of C to A is the complement to the G to T substitution in the *HBM* gene. This mutation causes a substitution of glycine 171 with valine (G171V).

The HBM polymorphism was confirmed by examining the DNA sequence of different groups of individuals. In all members of the HBM pedigree (38 individuals), the HBM polymorphism was observed in the heterozygous form in affected (i.e., elevated bone mass)

individuals only (N=18). In unaffected relatives (N=20) (BMDZ<2.0) the HBM polymorphism was never observed. To determine whether this gene was ever observed in individuals outside of the HBM pedigree, 297 phenotyped individuals were characterized at the site of the *HBM* gene. None were heterozygous at the site of the HBM polymorphism. In an unphenotyped control group, 1 of 42 individuals was observed to be heterozygous at position 582. Since this individual is deceased, their bone mineral density could not be obtained. Taken together, these data prove that the polymorphism observed in the kindred displaying the high bone mass phenotype is strongly correlated with the G→T polymorphism at position 582 of *Zmax1*. Taken together, these results establish that the HBM polymorphism genetically segregates with the HBM phenotype, and that both the HBM polymorphism and phenotype are rare in the general population.

XII. Allele Specific Oligonucleotide (ASO) Analysis

The amplicon containing the HBM1 polymorphism was PCR amplified using primers specific for the exon of interest. The appropriate population of individuals was PCR amplified in 96 well microtiter plates as follows. PCR reactions (20 µl) containing 1X Promega PCR buffer (Cat. # M1883 containing 1.5 mM MgCl₂), 100mM dNTP, 200 nM PCR primers (1863F: CCAAGTTCTGAGAAGTCC and 1864R: AATACCTGAAACCAT ACCTG), 1 U Amplitaq, and 20 ng of genomic DNA were prepared and amplified under the following PCR conditions: 94°C, 1 minute, (94°C, 30 sec.; 58°C, 30 sec.; 72°C, 1 min.) X35 cycles), 72°C, 5', 4°C, hold. Loading dye was then added and 10 µl of the products was electrophoresed on 1.5% agarose gels containing 1 µg/ml ethidium bromide at 100-150 V for 5-10 minutes. Gels were treated 20 minutes in denaturing solution (1.5 M NaCl, 0.5 N NaOH), and rinsed briefly with water. Gels were then neutralized in 1 M Tris-HCl, pH 7.5,

1.5 M NaCl, for 20 minutes and rinsed with water. Gels were soaked in 10 X SSC for 20 minutes and blotted onto nylon transfer membrane (Hybond N+- Amersham) in 10X SSC overnight. Filters were the rinsed in 6X SSC for 10 minutes and UV crosslinked.

The allele specific oligonucleotides (ASO) were designed with the polymorphism approximately in the middle. Oligonucleotides were phosphate free at the 5' end and were purchased from Gibco BRL. Sequences of the oligonucleotides are:

2326 Zmax1.ASO.g: AGACTGGGGTGAGACGC

2327 Zmax1.ASO.t: CAGACTGGGTTGAGACGCC

The polymorphic nucleotides are underlined. To label the oligos, 1.5 µl of 1 µg/µl ASO oligo (2326.Zmax1.ASO.g or 2327.Zmax1.ASO.t), 11 µl ddH₂O, 2 µl 10X kinase forward buffer, 5 µl γ-³²P-ATP (6000 Ci/mMole), and 1 µl T4 polynucleotide kinase (10 U/µl) were mixed, and the reaction incubated at 37°C for 30-60 minutes. Reactions were then placed at 95°C for 2 minutes and 30 ml H₂O was added. The probes were purified using a G25 microspin column (Pharmacia).

Blots were prehybridized in 10 ml 5X SSPE, 5X Denhardt's, 2% SDS, and 100 µg/ml, denatured, sonicated salmon sperm DNA at 40°C for 2 hr. The entire reaction mix of kinased oligo was then added to 10 ml fresh hybridization buffer (5X SSPE, 5X Denhardts, 2% SDS) and hybridized at 40°C for at least 4 hours to overnight.

All washes done in 5X SSPE, 0.1 % SDS. The first wash was at 45°C for 15 minutes; the solution was then changed and the filters washed 50°C for 15 minutes. Filters were then exposed to Kodak biomax film with 2 intensifying screens at -70°C for 15 minutes to 1 hr. If necessary the filters were washed at 55°C for 15 minutes and exposed to film again. Filters were stripped by washing in boiling 0.1X SSC, 0.1% SDS for 10 minutes at least 3 times.

The two films that best captured the allele specific assay with the 2 ASOs were converted into digital images by scanning them into Adobe PhotoShop. These images were overlaid against each other in Graphic Converter and then scored and stored in FileMaker Pro 4.0 (see Fig. 9).

5 XIII. Cellular Localization of *Zmax1*

A. *Gene Expression in Rat tibia by non isotopic In Situ Hybridization*

In situ hybridization was conducted by Pathology Associates International (PAI), Frederick, MD. This study was undertaken to determine the specific cell types that express the *Zmax1* gene in rat bone with particular emphasis on areas of bone growth and remodeling.

10 *Zmax1* probes used in this study were generated from both human (Hu*Zmax1*) and mouse (Ms*Zmax1*) cDNAs, which share an 87% sequence identity. The homology of human and mouse *Zmax1* with rat *Zmax1* is unknown.

For example, gene expression by non-isotopic *in situ* hybridization was performed as follows, but other methods would be known to the skilled artisan. Tibias were collected from

15 two 6 to 8 week old female Sprague Dawley rats euthanized by carbon dioxide asphyxiation. Distal ends were removed and proximal tibias were snap frozen in OCT embedding medium with liquid nitrogen immediately following death. Tissues were stored in a -80°C freezer.

Probes for amplifying PCR products from cDNA were prepared as follows. The primers to amplify PCR products from a cDNA clone were chosen using published sequences

20 of both human LRP5 (Genbank Accession No. ABO17498) and mouse LRP5 (Genbank Accession No. AFO64984). In order to minimize cross reactivity with other genes in the LDL receptor family, the PCR products were derived from an intracellular portion of the protein coding region. PCR was performed in a 50 µl reaction volume using cDNA clone as

template. PCR reactions contained 1.5 mM MgCl₂, 1 unit Amplitaq, 200 μM dNTPs and 2 μM each primer. PCR cycling conditions were 94°C for 1 min., followed by 35 cycles of 94°C for 30 seconds, 55°C for 30 seconds, 72°C for 30 seconds; followed by a 5 minute extension at 72°C. The reactions were then run on a 1.5 % agarose Tris-Acetate gel. DNA was eluted from the agarose, ethanol precipitated and resuspended in 10 mM Tris, pH 8.0. Gel purified PCR products were prepared for both mouse and human cDNAs and supplied to Pathology Associates International for *in situ* hybridizations.

The sequence of the human and mouse PCR primers and products were as follows:

Human Zmax1 sense primer (HBM1253)

10 CCGTGTGCTCCGCCGCCAGTTC

Human Zmax1 antisense primer (HBM1465)

GGCTCACGGAGCTCATCATGGACTT

Human Zmax1 PCR product

15 CCGTGTGCTCCGCCGCCAGTTCCTGCGCGGGGTGAGTGTGTGGACCTGCGCCTGCGCTGCGACGGCGAG
GCAGACTGTCAGGACCGCTCAGACGAGGTGGACTGTGACGCCATCTGCCTGCCCAACCAGTTCCGGTGTGCGAGC
GGCCAGTGTGCTCATCAAACAGCAGTGGACTCCTTCCCCGACTGTATCGACGGCTCCGACGAGCTCATGTGT
GAAATCACCAAGCCGCCCTCAGACGACAGCCCGGCCACAGCAGTGCCATCGGGCCCGTCATTGGCATCATCCTC
TCTCTCTTCGTCATGGGTGGTGTCTATTTTGTGTGCCAGCGGTGGTGTGCCAGCGCTATGCGGGGGCCAACGGG
CCCTTCCCGCACGAGTATGTCAGCGGGACCCCGCACGTGCCCCTCAATTTATAGCCCCGGGCGGTTCACGAT
20 GGCCCCCTTCACAGGCATCGCATGCGGAAAGTCCATGATGAGCTCCGTGAGCC

Mouse Zmax1 Sense primer (HBM1655)

AGCGAGGCCACCATCCACAGG

Mouse Zmax1 antisense primer (HBM1656)

TCGCTGGTCGGCATAATCAAT

Mouse Zmax1 PCR product

AGCAGAGCCACCATCCACAGGATCTCCCTGGAGACTAACAACAACGATGTGGCTATCCCACTCACGGGTGTCAAA
 GAGGCCTCTGCACTGGACTTTGATGTGTCCAACAATCACATCTACTGGACTGATGTTAGCCTCAAGACGATCAGC
 CGAGCCTTCATGAATGGGAGCTCAGTGGAGCACGTGATTGAGTTTGGCCTCGACTACCCTGAAGGAATGGCTGTG
 5 GACTGGATGGGCAAGAACCTCTATTGGGCGGACACAGGGACCAACAGGATTGAGGTGGCCCGGCTGGATGGGCAG
 TTCCGGCAGGTGCTTGTGTGGAGAGACCTTGACAACCCCAGGTCTCTGGCTCTGGATCCTACTAAAGGCTACATC
 TACTGGACTGAGTGGGGTGGCAAGCCAAGGATTGTGCGGGCCTTCATGGATGGGACCAATTGTATGACACTGGTA
 GACAAGGTGGGCGGGCCCAACGACCTCACCATTGATTATGCCGACCAGCGA

Riboprobes were synthesized as follows. The PCR products were reamplified with
 10 chimeric primers designed to incorporate either a T3 promoter upstream, or a T7 promoter
 downstream of the reamplification products. The resulting PCR products were used as
 template to synthesize digoxigenin-labeled riboprobes by *in vitro* transcription (IVT).
 Antisense and sense riboprobes were synthesized using T7 and T3 RNA polymerases,
 respectively, in the presence of digoxigenin-11-UTP (Boehringer-Mannheim) using a
 15 MAXIscript IVT kit (Ambion) according to the manufacturer. The DNA was then degraded
 with Dnase-1, and unincorporated digoxigenin was removed by ultrafiltration. Riboprobe
 integrity was assessed by electrophoresis through a denaturing polyacrylamide gel.
 Molecular size was compared with the electrophoretic mobility of a 100-1000 base pair (bp)
 RNA ladder (Ambion). Probe yield and labeling was evaluated by blot immunochemistry.
 20 Riboprobes were stored in 5 μ l aliquots at -80°C .

The *in situ* hybridization was performed as follows. Frozen rat bone was cut into 5
 μM sections on a Jung CM3000 cryostat (Leica) and mounted on adhesive slides
 (Instrumedics). Sections were kept in the cryostat at -20°C until all the slides were prepared
 in order to prevent mRNA degradation prior to post-fixation for 15 minutes in 4%
 25 paraformaldehyde. Following post-fixation, sections were incubated with 1 ng/ μ l of either

antisense or sense riboprobe in Pathology Associates International (PAI) customized hybridization buffer for approximately 40 hours at 58°C. Following hybridization, slides were subjected to a series of post-hybridization stringency washes to reduce nonspecific probe binding. Hybridization was visualized by immunohistochemistry with an anti-
5 digoxigenin antibody (FAB fragment) conjugated to alkaline phosphatase. Nitroblue tetrazolium chloride/bromochloroindolyl phosphate (Boehringer-Mannheim), a precipitating alkaline phosphatase substrate, was used as the chromogen to stain hybridizing cells purple to nearly black, depending on the degree of staining. Tissue sections were counter-stained with nuclear fast red. Assay controls included omission of the probe, omission of probe and anti-
10 digoxigenin antibody.

Specific cell types were assessed for demonstration of hybridization with antisense probes by visualizing a purple to black cytoplasmic and/or peri-nuclear staining indicating a positive hybridization signal for mRNA. Each cell type was compared to the replicate sections, which were hybridized with the respective sense probe. Results were considered
15 positive if staining was observed with the antisense probe and no staining or weak background with the sense probe.

The cellular localization of the hybridization signal for each of the study probes is summarized in Table 5. Hybridization for *Zmax1* was primarily detected in areas of bone involved in remodeling, including the endosteum and trabecular bone within the metaphysis.
20 Hybridization in selected bone lining cells of the periosteum and epiphysis were also observed. Positive signal was also noted in chondrocytes within the growth plate, particularly in the proliferating chondrocytes. See Figs. 10, 11 and 12 for representative photomicrographs of *in situ* hybridization results.

TABLE 5

Summary of Zmax1 *in situ* hybridization in rat tibia

PROBE	SITE	ISH SIGNAL
<u>Hu Zmax1</u>	<u>Epiphysis</u>	
	Osteoblasts	+
	Osteoclasts	-
	<u>Growth Plate</u>	
	resting chondrocytes	-
	proliferating chondrocytes	+
	hypertrophic chondrocytes	-
	<u>Metaphysis</u>	
	osteoblasts	+
	osteoclasts	+
	<u>Diaphysis</u>	-
	<u>Endosteum</u>	
	osteoblasts	+
	osteoclasts	+
	<u>Periosteum</u>	-
<u>MsZmax1</u>	<u>Epiphysis</u>	
	Osteoblasts	+
	Osteoclasts	-
	<u>Growth Plate</u>	
	resting chondrocytes	-
	proliferating chondrocytes	+
	hypertrophic chondrocytes	+
	<u>Metaphysis</u>	
	osteoblasts	+
	osteoclasts	+
	<u>Diaphysis</u>	-
	<u>Endosteum</u>	
	osteoblasts	+
	osteoclasts	+
	<u>Periosteum</u>	+

Legend: "+" = hybridization signal detected "-" = no hybridization signal detected

"ISH" - *In situ* hybridization

These studies confirm the positional expression of *Zmax1* in cells involved in bone remodeling and bone formation. *Zmax1* expression in the zone of proliferation and in the osteoblasts and osteoclasts of the proximal metaphysis, suggests that the *Zmax1* gene is involved in the process of bone growth and mineralization. The activity and differentiation of osteoblasts and osteoclasts are closely coordinated during development as bone is formed and during growth as well as in adult life as bone undergoes continuous remodeling. The formation of internal bone structures and bone remodeling result from the coupling of bone resorption by activated osteoclasts with subsequent deposition of new material by osteoblasts. *Zmax1* is related to the LDL receptor gene, and thus may be a receptor involved in mechanosensation and subsequent signaling in the process of bone remodeling. Therefore, changes in the level of expression of this gene could impact on the rate of remodeling and degree of mineralization of bone. Similar studies can be designed for *in situ* analysis of HBM or *Zmax1* in other cells or tissues.

XIV. Antisense

Antisense oligonucleotides are short synthetic nucleic acids that contain complementary base sequences to a targeted RNA. Hybridization of the RNA in living cells with the antisense oligonucleotide interferes with RNA function and ultimately blocks protein expression. Therefore, any gene for which the partial sequence is known can be targeted by an antisense oligonucleotide.

Antisense technology is becoming a widely used research tool and will play an increasingly important role in the validation and elucidation of therapeutic targets identified by genomic sequencing efforts.

Antisense technology was developed to inhibit gene expression by utilizing an oligonucleotide complementary to the mRNA that encodes the target gene. There are several possible mechanisms for the inhibitory effects of antisense oligonucleotides. Among them, degradation of mRNA by RNase H is considered to be the major mechanism of inhibition of protein function. This technique was originally used to elucidate the function of a target gene, but may also have therapeutic applications, provided it is designed carefully and properly.

An example of materials and methods for preparing antisense oligonucleotides can be performed as follows. Preliminary studies have been undertaken in collaboration with Sequiter (Natick, MA) using the antisense technology in the osteoblast-like murine cell line, MC3T3. These cells can be triggered to develop along the bone differentiation sequence. An initial proliferation period is characterized by minimal expression of differentiation markers and initial synthesis of collagenous extracellular matrix. Collagen matrix synthesis is required for subsequent induction of differentiation markers. Once the matrix synthesis begins, osteoblast marker genes are activated in a clear temporal sequence: alkaline phosphatase is induced at early times while bone sialoprotein and osteocalcin appear later in the differentiation process. This temporal sequence of gene expression is useful in monitoring the maturation and mineralization process. Matrix mineralization, which does not begin until several days after maturation has started, involves deposition of mineral on and within collagen fibrils deep within the matrix near the cell layer-culture plate interface. The collagen fibril-associated mineral formed by cultured osteoblasts resembles that found in woven bone in vivo and therefore is used frequently as a study reagent.

MC3T3 cells were transfected with antisense oligonucleotides for the first week of the differentiation, according to the manufacturer's specifications (U.S. Patent No. 5,849,902).

The oligonucleotides designed for Zmax1 are given below:

10875: AGUACAGCUUCUUGCCAACCCAGUC

10876: UCCUCCAGGUCGAUGGUCAGCCCAU

10877: GUCUGAGUCCGAGUUCAAAUCCAGG

- 5 Figure 13 shows the results of antisense inhibition of Zmax1 in MC3T3 cells. The three oligonucleotides shown above were transfected into MC3T3 and RNA was isolated according to standard procedures. Northern analysis clearly shows markedly lower steady state levels of the Zmax1 transcript while the control gene GAPDH remained unchanged. Thus, antisense technology using the primers described above allows for the study of the role of
- 10 Zmax1 expression on bone biology. Similar primers can be used to study Zmax1 expression and its ability to regulate lipid levels in an animal.

- The protein encoded by Zmax1 is related to the Low Density Lipoprotein receptor (LDL receptor). See, Goldstein *et al.*, *Ann. Rev. Cell Biology*, 1:1-39 (1985); Brown *et al.*, *Science*, 232:34-47 (1986). The LDL receptor is responsible for uptake of low density
- 15 lipoprotein, a lipid-protein aggregate that includes cholesterol. Individuals with a defect in the LDL receptor are deficient in cholesterol removal and tend to develop atherosclerosis. In addition, cells with a defective LDL receptor show increased production of cholesterol, in part because of altered feedback regulation of cholesterol synthetic enzymes and in part because of increased transcription of the genes for these enzymes. In some cell types,
- 20 cholesterol is a precursor for the formation of steroid hormones.

Thus, the LDL receptor may, directly or indirectly, function as a signal transduction protein and may regulate gene expression. Because Zmax1 is related to the LDL receptor, this protein may also be involved in signaling between cells in a way that affects bone remodeling as well as regulate lipid levels and therefore lipid-mediated diseases.

The glycine 171 amino acid is likely to be important for the function of Zmax1 because this amino acid is also found in the mouse homologue of Zmax1. The closely related LRP6 protein also contains glycine at the corresponding position (Brown *et al.*, *Biochemical and Biophysical Research Comm.*, 248:879-888 (1988)). Amino acids that are important in a protein's structure or function tend to be conserved between species, because natural selection prevents mutations with altered amino acids at important positions from arising.

In addition, the extracellular domain of Zmax1 contains four repeats consisting of five YWT motifs followed by an EFG motif. This 5YWT+EGF repeat is likely to form a distinct folded protein domain, as this repeat is also found in the LDL receptor and other LDL receptor-related proteins. The first three 5YWT+EGF repeats are very similar in their structure, while the fourth is highly divergent. Glycine 171 occurs in the central YWT motif of the first 5YWT+EGF repeat in Zmax1. The other two similar 5YWT+EGF repeats of Zmax1 also contain glycine at the corresponding position, as does the 5YWT+EGF repeat in the LDL receptor protein. However, only 17.6% of the amino acids are identical among the first three 5YWT+EGF repeats in Zmax1 and the single repeat in the LDL receptor. These observations indicate that glycine 171 is essential to the function of this repeat, and mutation of glycine 171 causes a functional alteration of Zmax1. The cDNA and peptide sequences are shown in Figs. 6A-6E. The critical base at nucleotide position 582 is indicated in bold and is underlined.

Northern blot analysis (Figs. 7A-B) reveals that Zmax1 is expressed in human bone tissue as well as numerous other tissues. A multiple-tissue Northern blot (Clontech, Palo Alto, CA) was probed with exons from Zmax1. As shown in Fig. 7A, the 5.5 kb Zmax1 transcript was highly expressed in heart, kidney, lung, liver and pancreas and is expressed at lower levels in skeletal muscle and brain. A second northern blot, shown in Fig. 7B,

confirmed the transcript size at 5.5 kb, and indicated that *Zmax1* is expressed in bone, bone marrow, calvaria and human osteoblastic cell lines.

Taken together, these results indicate that the HBM polymorphism in the *Zmax1* gene is responsible for the HBM phenotype, and that the *Zmax1* gene is important in bone
5 development. In addition, because mutation of *Zmax1* can alter bone mineralization and development as well as lipid levels, it is likely that molecules that bind to *Zmax1* may usefully alter bone development and lipid levels. Such molecules may include, for example, small molecules, proteins, RNA aptamers, peptide aptamers, and the like.

XV. Preparation of Nucleic Acids, Vectors, Transformations and Host Cells

10 Large amounts of the nucleic acids of the present invention may be produced by replication in a suitable host cell. Natural or synthetic nucleic acid fragments coding for a desired fragment will be incorporated into recombinant nucleic acid constructs, usually DNA constructs, capable of introduction into and replication in a prokaryotic or eukaryotic cell. Usually the nucleic acid constructs will be suitable for replication in a unicellular host, such
15 as yeast or bacteria, but may also be intended for introduction to (with and without integration within the genome) cultured mammalian or plant or other eukaryotic cell lines. The purification of nucleic acids produced by the methods of the present invention is described, for example, in Sambrook *et al.*, *Molecular Cloning. A Laboratory Manual*, 2nd Ed. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989) or Ausubel *et al.*,
20 *Current Protocols in Molecular Biology*, J. Wiley and Sons, NY (1992).

The nucleic acids of the present invention may also be produced by chemical synthesis, e.g., by the phosphoramidite method described by Beaucage *et al.*, *Tetra. Letts.*, 22:1859-1862 (1981) or the triester method according to Matteucci, *et al.*, *J. Am. Chem. Soc.*,

103:3185 (1981), and may be performed on commercial, automated oligonucleotide synthesizers. A double-stranded fragment may be obtained from the single-stranded product of chemical synthesis either by synthesizing the complementary strand and annealing the strands together under appropriate conditions or by adding the complementary strand using

- 5 DNA polymerase with an appropriate primer sequence.

Nucleic acid constructs prepared for introduction into a prokaryotic or eukaryotic host may comprise a replication system recognized by the host, including the intended nucleic acid fragment encoding the desired protein, and will preferably also include transcription and translational initiation regulatory sequences operably linked to the protein encoding segment.

- 10 Expression vectors may include, for example, an origin of replication or autonomously replicating sequence (ARS) and expression control sequences, a promoter, an enhancer and necessary processing information sites, such as ribosome-binding sites, RNA splice sites, polyadenylation sites, transcriptional terminator sequences, and mRNA stabilizing sequences. Secretion signals may also be included where appropriate, whether from a native HBM or
- 15 Zmax1 protein or from other receptors or from secreted proteins of the same or related species, which allow the protein to cross and/or lodge in cell membranes, and thus attain its functional topology, or be secreted from the cell. Such vectors may be prepared by means of standard recombinant techniques well known in the art and discussed, for example, in Sambrook *et al.*, *Molecular Cloning. A Laboratory Manual*, 2nd Ed. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989) or Ausubel *et al.*, *Current Protocols in Molecular Biology*, J. Wiley and Sons, NY (1992).
- 20

An appropriate promoter and other necessary vector sequences will be selected so as to be functional in the host, and may include, when appropriate, those naturally associated with Zmax1 or HBM genes. Examples of workable combinations of cell lines and expression

vectors are described in Sambrook *et al.*, *Molecular Cloning. A Laboratory Manual*, 2nd Ed. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989) or Ausubel *et al.*, *Current Protocols in Molecular Biology*, J. Wiley and Sons, NY (1992). Many useful vectors are known in the art and may be obtained from such vendors as Stratagene, New England BioLabs, Promega Biotech, and others. Promoters such as the *trp*, *lac* and phage promoters, tRNA promoters and glycolytic enzyme promoters may be used in prokaryotic hosts. Useful yeast promoters include promoter regions for metallothionein, 3-phosphoglycerate kinase or other glycolytic enzymes such as enolase or glyceraldehyde-3-phosphate dehydrogenase, enzymes responsible for maltose and galactose utilization, and others. Vectors and promoters suitable for use in yeast expression are further described in EP 73,675A. Appropriate non-native mammalian promoters might include the early and late promoters from SV40 (Fiers *et al.*, *Nature*, 273:113 (1978)) or promoters derived from murine Moloney leukemia virus, mouse tumor virus, avian sarcoma viruses, adenovirus II, bovine papilloma virus or polyoma. In addition, the construct may be joined to an amplifiable gene (e.g., DHFR) so that multiple copies of the gene may be made. For appropriate enhancer and other expression control sequences, see also *Enhancers and Eukaryotic Gene Expression*, Cold Spring Harbor Press, Cold Spring Harbor, NY (1983).

While such expression vectors may replicate autonomously, they may also replicate by being inserted into the genome of the host cell, by methods well known in the art.

Expression and cloning vectors will likely contain a selectable marker, a gene encoding a protein necessary for survival or growth of a host cell transformed with the vector. The presence of this gene ensures growth of only those host cells which express the inserts. Typical selection genes encode proteins that a) confer resistance to antibiotics or other toxic substances, e.g. ampicillin, neomycin, methotrexate, etc.; b) complement auxotrophic

iciencies, or c) supply critical nutrients not available from complex media, e.g., the gene encoding D-alanine racemase for Bacilli. The choice of the proper selectable marker will depend on the host cell, and appropriate markers for different hosts are well known in the art.

The vectors containing the nucleic acids of interest can be transcribed *in vitro*, and the
5 resulting RNA introduced into the host cell by well-known methods, e.g., by injection (see, Kubo *et al.*, *FEBS Letts.* 241:119 (1988)), or the vectors can be introduced directly into host cells by methods well known in the art, which vary depending on the type of cellular host, including electroporation; transfection employing calcium chloride, rubidium chloride, calcium phosphate, DEAE-dextran, or other substances; microprojectile bombardment;
10 lipofection; infection (where the vector is an infectious agent, such as a retroviral genome); and other methods. See generally, Sambrook *et al.*, 1989 and Ausubel *et al.*, 1992. The introduction of the nucleic acids into the host cell by any method known in the art, including those described above, will be referred to herein as "transformation." The cells into which have been introduced nucleic acids described above are meant to also include the progeny of
15 such cells.

Large quantities of the nucleic acids and proteins of the present invention may be prepared by expressing the Zmax1 or HBM nucleic acids or portions thereof in vectors or other expression vehicles in compatible prokaryotic or eukaryotic host cells. The most commonly used prokaryotic hosts are strains of *Escherichia coli*, although other prokaryotes,
20 such as *Bacillus subtilis* or *Pseudomonas* may also be used.

Mammalian or other eukaryotic host cells, such as those of yeast, filamentous fungi, plant, insect, or amphibian or avian species, may also be useful for production of the proteins of the present invention. Propagation of mammalian cells in culture is per se well known. See, Jakoby and Pastan (eds.), *Cell Culture. Methods in Enzymology*, volume 58, Academic

Press, Inc., Harcourt Brace Jovanovich, NY, (1979)). Examples of commonly used mammalian host cell lines are VERO and HeLa cells, Chinese hamster ovary (CHO) cells, and WI38, BHK, and COS cell lines, although it will be appreciated by the skilled practitioner that other cell lines may be appropriate, e.g., to provide higher expression

5 desirable glycosylation patterns, or other features.

Clones are selected by using markers depending on the mode of the vector construction. The marker may be on the same or a different DNA molecule, preferably the same DNA molecule. In prokaryotic hosts, the transformant may be selected, e.g., by resistance to ampicillin, tetracycline or other antibiotics. Production of a particular product
10 based on temperature sensitivity may also serve as an appropriate marker.

Prokaryotic or eukaryotic cells transformed with the nucleic acids of the present invention will be useful not only for the production of the nucleic acids and proteins of the present invention, but also, for example, in studying the characteristics of *Zmax1* or *HBM* proteins.

15 Antisense nucleic acid sequences are useful in preventing or diminishing the expression of *Zmax1* or *HBM*, as will be appreciated by one skilled in the art. For example, nucleic acid vectors containing all or a portion of the *Zmax1* or *HBM* gene or other sequences from the *Zmax1* or *HBM* region may be placed under the control of a promoter in an antisense orientation and introduced into a cell. Expression of such an antisense construct
20 within a cell will interfere with *Zmax1* or *HBM* transcription and/or translation and/or replication.

The probes and primers based on the *Zmax1* and *HBM* gene sequences disclosed herein are used to identify homologous *Zmax1* and *HBM* gene sequences and proteins in other species. These *Zmax1* and *HBM* gene sequences and proteins are used in the

diagnostic/prognostic, therapeutic and drug screening methods described herein for the species from which they have been isolated.

XVI. Protein Expression and Purification

Expression and purification of the HBM protein of the invention can be performed essentially as outlined below. To facilitate the cloning, expression and purification of membrane and secreted protein from the *HBM* gene, a gene expression system, such as the pET System (Novagen), for cloning and expression of recombinant proteins in *E. coli* was selected. Also, a DNA sequence encoding a peptide tag, the His-Tap, was fused to the 3' end of DNA sequences of interest to facilitate purification of the recombinant protein products. The 3' end was selected for fusion to avoid alteration of any 5' terminal signal sequence.

Nucleic acids chosen, for example, from the nucleic acids set forth in SEQ ID NOS: 1, 3 and 5-12 for cloning HBM were prepared by polymerase chain reaction (PCR). Synthetic oligonucleotide primers specific for the 5' and 3' ends of the HBM nucleotide sequence were designed and purchased from Life Technologies (Gaithersburg, MD). All forward primers (specific for the 5' end of the sequence) were designed to include an NcoI cloning site at the 5' terminus. These primers were designed to permit initiation of protein translation at the methionine residue encoded within the NcoI site followed by a valine residue and the protein encoded by the HBM DNA sequence. All reverse primers (specific for the 3' end of the sequence) included an EcoRI site at the 5' terminus to permit cloning of the HBM sequence into the reading frame of the pET-28b. The pET-28b vector provided a sequence encoding an additional 20 carboxyl-terminal amino acids including six histidine residues (at the C-terminus), which comprised the histidine affinity tag.

Genomic DNA prepared from the *HBM* gene was used as the source of template DNA for PCR amplification (Ausubel *et al.*, *Current Protocols in Molecular Biology*, John Wiley & Sons (1994)). To amplify a DNA sequence containing the HBM nucleotide sequence, genomic DNA (50 ng) was introduced into a reaction vial containing 2 mM MgCl₂, 1 μM
5 synthetic oligonucleotide primers (forward and reverse primers) complementary to and flanking a defined HBM, 0.2 mM of each of deoxynucleotide triphosphate, dATP, dGTP, dCTP, dTTP and 2.5 units of heat stable DNA polymerase (Amplitaq, Roche Molecular Systems, Inc., Branchburg, NJ) in a final volume of 100 μl.

Upon completion of thermal cycling reactions, each sample of amplified DNA was
10 purified using the Qiaquick Spin PCR purification kit (Qiagen, Gaithersburg, MD). All amplified DNA samples were subjected to digestion with the restriction endonucleases, e.g., NcoI and EcoRI (New England BioLabs, Beverly, MA) (Ausubel *et al.*, *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc. (1994)). DNA samples were then subjected to electrophoresis on 1.0% NuSeive (FMC BioProducts, Rockland, ME) agarose gels. DNA
15 was visualized by exposure to ethidium bromide and long wave UV irradiation. DNA contained in slices isolated from the agarose gel was purified using the Bio 101 GeneClean Kit protocol (Bio 101, Vista, CA).

The pET-28b vector was prepared for cloning by digestion with restriction endonucleases, e.g., NcoI and EcoRI (New England BioLabs, Beverly, MA) (Ausubel *et al.*,
20 *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc. (1994)). The pET-28a vector, which encodes the histidine affinity tag that can be fused to the 5' end of an inserted gene, was prepared by digestion with appropriate restriction endonucleases.

Following digestion, DNA inserts were cloned (Ausubel *et al.*, *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc. (1994)) into the previously digested pET-28b

expression vector. Products of the ligation reaction were then used to transform the BL21 strain of *E. coli* (Ausubel *et al.*, *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc. (1994)) as described below.

Competent bacteria, *E. coli* strain BL21 or *E. coli* strain BL21 (DE3), were
5 transformed with recombinant pET expression plasmids carrying the cloned HBM sequence according to standard methods (Ausubel *et al.*, *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc. (1994)). Briefly, 1 μ l of ligation reaction was mixed with 50 μ l of electrocompetent cells and subjected to a high voltage pulse, after which samples were incubated in 0.45 ml SOC medium (0.5% yeast extract, 2.0% tryptone, 10 mM NaCl, 2.5 mM
10 KCl, 10 mM MgCl₂, 10 mM MgSO₄ and 20 mM glucose) at 37°C with shaking for 1 hour. Samples were then spread on LB agar plates containing 25 μ g/ml kanamycin sulfate for growth overnight. Transformed colonies of BL21 were then picked and analyzed to evaluate cloned inserts, as described below.

Individual BL21 clones transformed with recombinant pET-28b HBM nucleotide
15 sequences were analyzed by PCR amplification of the cloned inserts using the same forward and reverse primers specific for the HBM sequences that were used in the original PCR amplification cloning reactions. Successful amplification verifies the integration of the HBM sequence in the expression vector (Ausubel *et al.*, *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc. (1994)).

20 Individual clones of recombinant pET-28b vectors carrying properly cloned HBM nucleotide sequences were picked and incubated in 5 ml of LB broth plus 25 μ g/ml kanamycin sulfate overnight. The following day plasmid DNA was isolated and purified using the Qiagen plasmid purification protocol (Qiagen Inc., Chatsworth, CA).

The pET vector can be propagated in any *E. coli* K-12 strain, e.g., HMS174, HB101, JM109, DH5 and the like, for purposes of cloning or plasmid preparation. Hosts for expression include *E. coli* strains containing a chromosomal copy of the gene for T7 RNA polymerase. These hosts were lysogens of bacteriophage DE3, a lambda derivative that carries the lacI gene, the lacUV5 promoter and the gene for T7 RNA polymerase. T7 RNA polymerase was induced by addition of isopropyl- β -D-thiogalactoside (IPTG), and the T7 RNA polymerase transcribes any target plasmid containing a functional T7 promoter, such as pET-28b, carrying its gene of interest. Strains include, for example, BL21(DE3) (Studier *et al.*, *Meth. Enzymol.*, 185:60-89 (1990)).

10 To express the recombinant HBM sequence, 50 ng of plasmid DNA are isolated as described above to transform competent BL21(DE3) bacteria as described above (provided by Novagen as part of the pET expression kit). The lacZ gene (β -galactosidase) is expressed in the pET-System as described for the HBM recombinant constructions. Transformed cells were cultured in SOC medium for 1 hour, and the culture was then plated on LB plates containing 25 μ g/ml kanamycin sulfate. The following day, the bacterial colonies were

15 pooled and grown in LB medium containing kanamycin sulfate (25 μ g/ml) to an optical density at 600 nm of 0.5 to 1.0 O.D. units, at which point 1 mM IPTG was added to the culture for 3 hours to induce gene expression of the HBM recombinant DNA constructions.

After induction of gene expression with IPTG, bacteria were collected by

20 centrifugation in a Sorvall RC-3B centrifuge at 3500 x g for 15 minutes at 4°C. Pellets were resuspended in 50 ml of cold mM Tris-HCl, pH 8.0; 0.1 M NaCl and 0.1 mM EDTA (STE buffer). Cells were then centrifuged at 2000 x g for 20 minutes at 4°C. Wet pellets were weighed and frozen at -80°C until ready for protein purification.

A variety of methodologies known in the art can be used to purify the isolated proteins (Coligan *et al.*, *Current Protocols in Protein Science*, John Wiley & Sons (1995)). For example, the frozen cells can be thawed, resuspended in buffer and ruptured by several passages through a small volume microfluidizer (Model M-110S, Microfluidics International Corp., Newton, MA). The resultant homogenate is centrifuged to yield a clear supernatant (crude extract) and, following filtration, the crude extract is fractioned over columns. Fractions are monitored by absorbance at OD₂₈₀ nm and peak fractions may be analyzed by SDS-PAGE.

The concentrations of purified protein preparations are quantified spectrophotometrically using absorbance coefficients calculated from amino acid content (Perkins, *Eur. J. Biochem.*, 157:169-180 (1986)). Protein concentrations are also measured by the method of Bradford, *Anal. Biochem.*, 72:248-254 (1976) and Lowry *et al.*, *J. Biol. Chem.*, 193:265-275 (1951) using bovine serum albumin as a standard.

SDS-polyacrylamide gels of various concentrations were purchased from BioRad (Hercules, CA), and stained with Coomassie blue. Molecular weight markers may include rabbit skeletal muscle myosin (200 kDa), *E. coli* β -galactosidase (116 kDa), rabbit muscle phosphorylase B (97.4 kDa), bovine serum albumin (66.2 kDa), ovalbumin (45 kDa), bovine carbonic anhydrase (31 kDa), soybean trypsin inhibitor (21.5 kDa), egg white lysozyme (14.4 kDa) and bovine aprotinin (6.5 kDa).

Once a sufficient quantity of the desired protein has been obtained, it may be used for various purposes. A typical use is the production of antibodies specific for binding. These antibodies may be either polyclonal or monoclonal, and may be produced by *in vitro* or *in vivo* techniques well known in the art. Monoclonal antibodies to epitopes of any of the peptides identified and isolated as described can be prepared from murine hybridomas

(Kohler, *Nature*, 256:495 (1975)). In summary, a mouse is inoculated with a few micrograms of HBM protein over a period of two weeks. The mouse is then sacrificed. The cells that produce antibodies are then removed from the mouse's spleen. The spleen cells are then fused with polyethylene glycol with mouse myeloma cells. The successfully fused cells are diluted in a microtiter plate and growth of the culture is continued. The amount of antibody per well is measured by immunoassay methods such as ELISA (Engvall, *Meth. Enzymol.*, 70:419 (1980)). Clones producing antibody can be expanded and further propagated to produce HBM antibodies. Other suitable techniques involve *in vitro* exposure of lymphocytes to the antigenic polypeptides, or alternatively, to selection of libraries of antibodies in phage or similar vectors. See Huse *et al.*, *Science*, 246:1275-1281 (1989). For additional information on antibody production see Davis *et al.*, *Basic Methods in Molecular Biology*, Elsevier, NY, Section 21-2 (1989).

Standard protocols for assessing the influence of an agent (*e.g.*, antibody, HBM protein, protein polymorphism or Zmax1 protein or compound) to alter lipid levels in a cell or the physiological levels in a subject are known. For example, see F.W. HEMMING, *LIPID ANALYSIS* (Bios Scientific Pub. 1996) and J. M. ORDOVAS, *LIPOPROTEIN PROTOCOLS* (Humana Press Inc. 1997). More specifically, cholesterol and triglyceride analysis can be performed using the Olympus AU5000 Cholesterol method. This method of measuring cholesterol combines the use of the enzymes with a modification of the peroxidase-phenol-4-aminoantipyrine system, substituting 2-hydroxy-3,5-dichlorobenzene sulfonic acid (2-OH 3,5 DCBSA) for the phenolic group for the measurement of total cholesterol in the subject serum. The assay is based on a series of coupled enzymatic reactions. Cholesterol esters present in serum are hydrolyzed to free cholesterol and fatty acids by cholesterol esterase. The cholesterol is in turn oxidized by cholesterol oxidase to cholest-4-en-3-one with the

simultaneous production of hydrogen peroxidase. The hydrogen peroxidase reacts with 4-aminoantipyrine in the presence of 2-OH-3,5-DCBSA to produce a chromophore that absorbs at 570 nm. The absorbance of the reaction mixture is measured biochromatically at 570/750 nm and is proportional to the cholesterol concentration of the sample.

- 5 For serum triglyceride analysis, the Olympus AU5000 triglyceride procedure can also be used. Briefly, it is based on a series of coupled enzymatic reactions. Triglycerides in the serum are hydrolyzed to free fatty acids and glycerol by lipoprotein lipase. Glycerol is phosphorylated enzymatically and then oxidized with glycerol phosphate-oxidase. The hydrogen peroxidase reacts with the chromogen 4-amino-antipyrine in the presence of DCB
- 10 Sulfonic Acid to give a chromophore with absorption which is measured bichromatically at 520/660 nm. The increase in absorbance of the reaction mixture is proportional to the triglyceride concentration of the sample.

XVII. Methods of Use: Gene Therapy

- In recent years, significant technological advances have been made in the area of gene
- 15 therapy for both genetic and acquired diseases. (Kay *et al.*, *Proc. Natl. Acad. Sci. USA*, 94:12744-12746 (1997)) Gene therapy can be defined as the deliberate transfer of DNA for therapeutic purposes. Improvement in gene transfer methods has allowed for development of gene therapy protocols for the treatment of diverse types of diseases. Gene therapy has also taken advantage of recent advances in the identification of new therapeutic genes,
- 20 improvement in both viral and nonviral gene delivery systems, better understanding of gene regulation, and improvement in cell isolation and transplantation.

The experiments below identify the *HBM* gene as a dominant mutation conferring elevated bone mass and that alters lipid levels. The fact that this mutation is dominant

indicates that expression of the HBM protein causes elevated bone mass and perhaps changes in lipid levels. Older individuals carrying the *HBM* gene, and, therefore expressing the HBM protein, do not suffer from osteoporosis. These individuals are equivalent to individuals being treated with the HBM protein. These observations are a strong experimental indication
5 that therapeutic treatment with the HBM protein prevents osteoporosis. The bone mass elevating activity of the *HBM* gene is termed "HBM function."

Therefore, according to the present invention, a method is also provided of supplying HBM function to mesenchymal stem cells (Onyia *et al.*, *J. Bone Miner. Res.*, 13:20-30 (1998); Ko *et al.*, *Cancer Res.*, 56:4614-4619 (1996)). Supplying such a function provides
10 protection against osteoporosis. For regulating lipid levels, HBM function can be supplied to liver cells, as well as other cells involved in lipid metabolism and lipid regulation (e.g., muscle cells, lesion cells, lipid laden foam cells and megakaryoblasts). The *HBM* gene or a part of the gene may be introduced into the cell in a vector such that the gene remains extrachromosomal. In such a situation, the gene will be expressed by the cell from the
15 extrachromosomal location.

Vectors for introduction of genes both for recombination and for extrachromosomal maintenance are known in the art, and any suitable vector may be used. Methods for
introducing DNA into cells such as electroporation, calcium phosphate co-precipitation, and viral transduction are known in the art, and the choice of method is within the competence of
20 one skilled in the art (Robbins, Ed., *Gene Therapy Protocols*, Human Press, NJ (1997)). Cells transformed with the *HBM* gene can be used as model systems to study osteoporosis and drug treatments that promote bone growth as well as to study lipid-mediated diseases.

As generally discussed above, the *HBM* gene or fragment, where applicable, may be used in gene therapy methods in order to increase the amount of the expression products of

such genes in mesenchymal stem cells or in other cells. It may be useful also to increase the level of expression of a given HBM protein, or a fragment thereof, even in those cells in which the wild type gene is expressed normally. Gene therapy would be carried out according to generally accepted methods as described by, for example, Friedman, *Therapy for*

5 *Genetic Diseases*, Friedman, Ed., Oxford University Press, pages 105-121 (1991).

A virus or plasmid vector containing a copy of the *HBM* gene linked to expression control elements and capable of replicating inside mesenchymal stem cells or liver cells, is prepared. Suitable vectors are known and described, for example, in U.S. Patent No. 5,252,479 and WO 93/07282, the disclosures of which are incorporated by reference herein in

10 their entirety. The vector is then injected into the patient, either locally into the bone marrow or liver, or systemically (in order to reach any mesenchymal stem cells located at other sites, i.e., in the blood). If the transfected gene is not permanently incorporated into the genome of each of the targeted cells, the treatment may have to be repeated periodically.

Gene transfer systems known in the art may be useful in the practice of the gene

15 therapy methods of the present invention. These include viral and non-viral transfer methods. A number of viruses have been used as gene transfer vectors, including polyoma, i.e., SV40 (Madzak *et al.*, *J. Gen. Virol.*, 73:1533-1536 (1992)), adenovirus (Berkner, *Curr. Top. Microbiol. Immunol.*, 158:39-61 (1992); Berkner *et al.*, *Bio Techniques*, 6:616-629 (1988); Gorziglia *et al.*, *J. Virol.*, 66:4407-4412 (1992); Quantin *et al.*, *Proc. Natl. Acad. Sci. USA*,

20 89:2581-2584 (1992); Rosenfeld *et al.*, *Cell*, 68:143-155 (1992); Wilkinson *et al.*, *Nucl. Acids Res.*, 20:2233-2239 (1992); Stratford-Perricaudet *et al.*, *Hum. Gene Ther.*, 1:241-256 (1990)), vaccinia virus (Mackett *et al.*, *Biotechnology*, 24:495-499 (1992)), adeno-associated virus (Muzyczka, *Curr. Top. Microbiol. Immunol.*, 158:91-123 (1992); Ohi *et al.*, *Gene*, 89:279-282 (1990)), herpes viruses including HSV and EBV (Margolskee, *Curr. Top.*

- Microbiol. Immunol.*, 158:67-90 (1992); Johnson *et al.*, *J. Virol.*, 66:2952-2965 (1992); Fink *et al.*, *Hum. Gene Ther.*, 3:11-19 (1992); Breakfield *et al.*, *Mol. Neurobiol.*, 1:337-371 (1987); Fresse *et al.*, *Biochem. Pharmacol.*, 40:2189-2199 (1990)), and retroviruses of avian (Brandyopadhyay *et al.*, *Mol. Cell Biol.*, 4:749-754 (1984); Petropoulos *et al.*, *J. Virol.*, 66:3391-3397 (1992)), murine (Miller, *Curr. Top. Microbiol. Immunol.*, 158:1-24 (1992); Miller *et al.*, *Mol. Cell Biol.*, 5:431-437 (1985); Sorge *et al.*, *Mol. Cell Biol.*, 4:1730-1737 (1984); Mann *et al.*, *J. Virol.*, 54:401-407 (1985)), and human origin (Page *et al.*, *J. Virol.*, 64:5370-5276 (1990); Buchschalcher *et al.*, *J. Virol.*, 66:2731-2739 (1992)). Most human gene therapy protocols have been based on disabled murine retroviruses.
- 10 Non-viral gene transfer methods known in the art include chemical techniques such as calcium phosphate coprecipitation (Graham *et al.*, *Virology*, 52:456-467 (1973); Pellicer *et al.*, *Science*, 209:1414-1422 (1980)), mechanical techniques, for example microinjection (Anderson *et al.*, *Proc. Natl. Acad. Sci. USA*, 77:5399-5403 (1980); Gordon *et al.*, *Proc. Natl. Acad. Sci. USA*, 77:7380-7384 (1980); Brinster *et al.*, *Cell*, 27:223-231 (1981); Constantini *et al.*, *Nature*, 294:92-94 (1981)), membrane fusion-mediated transfer via liposomes (Felgner *et al.*, *Proc. Natl. Acad. Sci. USA*, 84:7413-7417 (1987); Wang *et al.*, *Biochemistry*, 28:9508-9514 (1989); Kaneda *et al.*, *J. Biol. Chem.*, 264:12126-12129 (1989); Stewart *et al.*, *Hum. Gene Ther.*, 3:267-275 (1992); Nabel *et al.*, *Science*, 249:1285-1288 (1990); Lim *et al.*, *Circulation*, 83:2007-2011 (1992)), and direct DNA uptake and receptor-mediated DNA
- 15 transfer (Wolff *et al.*, *Science*, 247:1465-1468 (1990); Wu *et al.*, *BioTechniques*, 11:474-485 (1991); Zenke *et al.*, *Proc. Natl. Acad. Sci. USA*, 87:3655-3659 (1990); Wu *et al.*, *J. Biol. Chem.*, 264:16985-16987 (1989); Wolff *et al.*, *BioTechniques*, 11:474-485 (1991); Wagner *et al.*, 1990; Wagner *et al.*, *Proc. Natl. Acad. Sci. USA*, 88:4255-4259 (1991); Cotten *et al.*, *Proc. Natl. Acad. Sci. USA*, 87:4033-4037 (1990); Curiel *et al.*, *Proc. Natl. Acad. Sci. USA*,
- 20

88:8850-8854 (1991); Curiel *et al.*, *Hum. Gene Ther.*, 3:147-154 (1991)). Viral-mediated gene transfer can be combined with direct *in vivo* vectors to the mesenchymal stem cells and not into the surrounding cells (Romano *et al.*, *In Vivo*, 12(1):59-67 (1998); Gonez *et al.*, *Hum. Mol. Genetics*, 7(12):1913-9 (1998)). Alternatively, the retroviral vector producer cell
5 line can be injected into the bone marrow (Culver *et al.*, *Science*, 256:1550-1552 (1992)). Injection of producer cells would then provide a continuous source of vector particles. This technique has been approved for use in humans with inoperable brain tumors.

In an approach which combines biological and physical gene transfer methods, plasmid DNA of any size is combined with a polylysine-conjugated antibody specific to the
10 adenovirus hexon protein, and the resulting complex is bound to an adenovirus vector. The trimolecular complex is then used to infect cells. The adenovirus vector permits efficient binding, internalization, and degradation of the endosome before the coupled DNA is damaged.

Liposome/DNA complexes have been shown to be capable of mediating direct *in vivo*
15 gene transfer. While in standard liposome preparations the gene transfer process is non-specific, localized *in vivo* uptake and expression have been reported in tumor deposits, for example, following direct *in situ* administration (Nabel, *Hum. Gene Ther.*, 3:399-410 (1992)).

XVIII. Methods of Use: Transformed Hosts, Development of Pharmaceuticals and Research Tools

20 Cells and animals that carry the *HBM* gene can be used as model systems to study and test for substances that have potential as therapeutic agents (Onyia *et al.*, *J. Bone Miner. Res.*, 13:20-30 (1998); Broder *et al.*, *Bone*, 21:225-235 (1997)). The cells are typically cultured mesenchymal stem cells or liver cells. These may be isolated from individuals with somatic

or germline *HBM* genes. Alternatively, the cell line can be engineered to carry the *HBM* gene, as described above. After a test substance is applied to the cells, the transformed phenotype of the cell is determined. Any trait of transformed cells can be assessed, including formation of bone matrix in culture (Broder *et al.*, *Bone*, 21:225-235 (1997)), mechanical properties (Kizer *et al.*, *Proc. Natl. Acad. Sci. USA*, 94:1013-1018 (1997)), and response to application of putative therapeutic agents.

Animals for testing therapeutic agents can be selected after treatment of germline cells or zygotes. Such treatments include insertion of the *Zmax1* gene, as well as insertion of the *HBM* gene and disrupted homologous genes. Alternatively, the inserted *Zmax1* gene(s) and/or *HBM* gene(s) of the animals may be disrupted by insertion or deletion mutation of other genetic alterations using conventional techniques, such as those described by, for example, Capecchi, *Science*, 244:1288 (1989); Valancuis *et al.*, *Mol. Cell Biol.*, 11:1402 (1991); Hasty *et al.*, *Nature*, 350:243 (1991); Shinkai *et al.*, *Cell*, 68:855 (1992); Mombaerts *et al.*, *Cell*, 68:869 (1992); Philpott *et al.*, *Science*, 256:1448 (1992); Snouwaert *et al.*, *Science*, 257:1083 (1992); Donehower *et al.*, *Nature*, 356:215 (1992). After test substances have been administered to the animals, the growth of bone or modulation of lipids must be assessed. If the test substance enhances the growth of bone or regulates lipid levels, then the test substance is a candidate therapeutic agent. These animal models provide an extremely important vehicle for potential therapeutic products. Preferred models for studying lipid modulation include mice (Smith *et al.*, *J. Intern. Med.*, 242: 99-109 (1997)) and guinea pigs.

Individuals carrying the *HBM* gene have elevated bone mass and altered lipid levels as discussed in the example below. The *HBM* gene causes this phenotype by altering the activities, levels, expression patterns, and modification states of other molecules involved in bone development. Using a variety of established techniques, it is possible to identify

molecules, preferably proteins or mRNAs, whose activities, levels, expression patterns, and modification states are different between systems containing the *Zmax 1* gene and systems containing the *HBM* gene. Such systems can be, for example, cell-free extracts, cells, tissues or living organisms, such as mice or humans. For a mutant form of *Zmax1*, a complete
5 deletion of *Zmax1*, mutations lacking the extracellular or intracellular portion of the protein, or any other mutation in the *Zmax1* gene may be used. It is also possible to use expression of antisense *Zmax1* RNA or oligonucleotides to inhibit production of the *Zmax1* protein. For a mutant form of *HBM*, a complete deletion of *HBM*, mutations lacking the extracellular or intracellular portion of the *HBM* protein, or any other mutation in the *HBM* gene may be
10 used. It is also possible to use expression of antisense *HBM* RNA or oligonucleotides to inhibit production of the *HBM* protein.

Molecules identified by comparison of *Zmax1* systems and *HBM* systems can be used as surrogate markers in pharmaceutical development or in diagnosis of human or animal bone disease. Alternatively, such molecules may be used in treatment of bone disease. *See*,
15 Schena *et al.*, *Science*, 270:467-470 (1995).

For example, a transgenic mouse carrying the *HBM* gene in the mouse homologue is constructed. A mouse of the genotype *HBM*/+ is viable, healthy and has elevated bone mass. To identify surrogate markers for elevated bone mass, *HBM*/+ (i.e., heterozygous) and isogenic +/+ (i.e., wild-type) mice are sacrificed. Bone tissue mRNA is extracted from each
20 animal, and a "gene chip" corresponding to mRNAs expressed in the +/+ individual is constructed. mRNA from different tissues is isolated from animals of each genotype, reverse-transcribed, fluorescently labeled, and then hybridized to gene fragments affixed to a solid support. The ratio of fluorescent intensity between the two populations is indicative of the relative abundance of the specific mRNAs in the +/+ and *HBM*/+ animals. Genes

encoding mRNAs over- and under-expressed relative to the wild-type control are candidates for genes coordinately regulated by the *HBM* gene. This strategy can be similarly used to study lipid regulation.

Mice also serve as the most common experimental animal model for atherosclerosis research. There are at least three ways of inducing atherosclerosis in mice: (1) diet induced, apoE deficiency-induced and LDL receptor-deficiency induced. The methods for using a mouse model for testing agents which modulate lipid levels *in vivo* can be performed as described in Smith *et al.*, *J. Intern. Med.* 242: 99-109 (1997).

One standard procedure for identification of new proteins that are part of the same signaling cascade as an already-discovered protein is as follows. Cells are treated with radioactive phosphorous, and the already-discovered protein is manipulated to be more or less active. The phosphorylation state of other proteins in the cell is then monitored by polyacrylamide gel electrophoresis and autoradiography, or similar techniques. Levels of activity of the known protein may be manipulated by many methods, including, for example, comparing wild-type mutant proteins using specific inhibitors such as drugs or antibodies, simply adding or not adding a known extracellular protein, or using antisense inhibition of the expression of the known protein (Tamura *et al.*, *Science*, 280(5369):1614-7 (1998); Meng, *EMBO J.*, 17(15):4391-403 (1998); Cooper *et al.*, *Cell*, 1:263-73 (1982)).

In another example, proteins with different levels of phosphorylation are identified in TE85 osteosarcoma cells expressing either a sense or antisense cDNA for Zmax1. TE85 cells normally express high levels of Zmax1 (Dong *et al.*, *Biochem. & Biophys. Res. Comm.*, 251:784-790 (1998)). Cells containing the sense construct express even higher levels of Zmax1, while cells expressing the antisense construct express lower levels. Cells are grown in the presence of ³²P, harvested, lysed, and the lysates run on SDS polyacrylamide gels to

separate proteins, and the gels subjected to autoradiography (Ausubel *et al.*, *Current Protocols in Molecular Biology*, John Wiley & Sons (1997)). Bands that differ in intensity between the sense and antisense cell lines represent phosphoproteins whose phosphorylation state or absolute level varies in response to levels of Zmax1. As an alternative to the ³²P-labeling, unlabeled proteins may be separated by SDS-PAGE and subjected to immunoblotting, using the commercially available anti-phosphotyrosine antibody as a probe (Thomas *et al.*, *Nature*, 376(6537):267-71 (1995)). As an alternative to the expression of antisense RNA, transfection with chemically modified antisense oligonucleotides can be used (Woelf *et al.*, *Nucleic Acids Res.*, 18(7):1763-9 (1990)).

Many bone disorders, such as osteoporosis, have a slow onset and a slow response to treatment. It is therefore useful to develop surrogate markers for bone development and mineralization. Such markers can be useful in developing treatments for bone disorders, and for diagnosing patients who may be at risk for later development of bone disorders. Examples of preferred markers are N- and C-terminal telopeptide markers described, for example, in U.S. Patent Nos. 5,455,179, 5,641,837 and 5,652,112, the disclosures of which are incorporated by reference herein in their entirety. In the area of HIV disease, CD4 counts and viral load are useful surrogate markers for disease progression (Vlahov *et al.*, *JAMA*, 279(1):35-40 (1998)). There is a need for analogous surrogate markers in the area of bone disease.

A surrogate marker can be any characteristic that is easily tested and relatively insensitive to non-specific influences. For example, a surrogate marker can be a molecule such as a protein or mRNA in a tissue or in blood serum. Alternatively, a surrogate marker may be a diagnostic sign such as sensitivity to pain, a reflex response or the like.

In yet another example, surrogate markers for elevated bone mass are identified using a pedigree of humans carrying the *HBM* gene. Blood samples are withdrawn from three individuals that carry the *HBM* gene, and from three closely related individuals that do not. Proteins in the serum from these individuals are electrophoresed on a two dimensional gel system, in which one dimension separates proteins by size, and another dimension separates proteins by isoelectric point (Epstein *et al.*, *Electrophoresis*, 17(11):1655-70 (1996)). Spots corresponding to proteins are identified. A few spots are expected to be present in different amounts or in slightly different positions for the HBM individuals compared to their normal relatives. These spots correspond to proteins that are candidate surrogate markers. The identities of the proteins are determined by microsequencing, and antibodies to the proteins can be produced by standard methods for use in diagnostic testing procedures. Diagnostic assays for HBM proteins or other candidate surrogate markers include using antibodies described in this invention and a reporter molecule to detect HBM in human body fluids, membranes, bones, cells, tissues or extracts thereof. The antibodies can be labeled by joining them covalently or noncovalently with a substance that provides a detectable signal. In many scientific and patent literature, a variety of reporter molecules or labels are described including radionuclides, enzymes, fluorescent, chemi-luminescent or chromogenic agents (U.S. Patent Nos. 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149; and 4,366,241).

Using these antibodies, the levels of candidate surrogate markers are measured in normal individuals and in patients suffering from a bone disorder, such as osteoporosis, osteoporosis pseudoglioma, Engelmann's disease, Ribbing's disease, hyperphosphatasemia, Van Buchem's disease, melorheostosis, osteopetrosis, pycnodysostosis, sclerosteosis, osteopoikilosis, acromegaly, Paget's disease, fibrous dysplasia, tubular stenosis, osteogenesis

imperfecta, hypoparathyroidism, pseudohypoparathyroidism, pseudopseudohypoparathyroidism, primary and secondary hyperparathyroidism and associated syndromes, hypercalciuria, medullary carcinoma of the thyroid gland, osteomalacia and other diseases including lipid-mediated diseases. Techniques for measuring levels of protein in serum in a clinical setting using antibodies are well established. A protein that is consistently present in higher or lower levels in individuals carrying a particular disease or type of disease is a useful surrogate marker.

A surrogate marker can be used in diagnosis of a bone disorder. For example, consider a child that present to a physician with a high frequency of bone fracture. The underlying cause may be child abuse, inappropriate behavior by the child, or a bone disorder. To rapidly test for a bone disorder, the levels of the surrogate marker protein are measured using the antibody described above.

Levels of modification states of surrogate markers can be measured as indicators of the likely effectiveness of a drug that is being developed. It is especially convenient to use surrogate markers in creating treatments for bone disorders, because alterations in bone development or mineralization may require a long time to be observed. For example, a set of bone mRNAs, termed the "HBM-inducible mRNA set" is found to be overexpressed in HBM/+ mice as compared to +/+ mice, as described above. Expression of this set can be used as a surrogate marker. Specifically, if treatment of +/+ mice with a compound results in overexpression of the HBM-inducible mRNA set, then that compound is considered a promising candidate for further development.

This invention is particularly useful for screening compounds by using the Zmax1 or HBM protein or binding fragment thereof in any of a variety of drug screening techniques.

The Zmax1 or HBM protein or fragment employed in such a test may either be free in solution, affixed to a solid support, or borne on a cell surface. One method of drug screening utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the protein or fragment, preferably in competitive binding assays.

5 Such cells, either in viable or fixed form, can be used for standard binding assays. One may measure, for example, for the formation of complexes between a Zmax1 or HBM protein or fragment and the agent being tested, or examine the degree to which the formation of a complex between a Zmax1 or HBM protein or fragment and a known ligand is interfered with by the agent being tested.

10 Thus, the present invention provides methods of screening for drugs comprising contacting such an agent with a Zmax1 or HBM protein, or fragment thereof, and assaying (i) for the presence of a complex between the agent and the Zmax1 or HBM protein or fragment, or (ii) for the presence of a complex between the Zmax1 or HBM protein or fragment and a ligand, by methods well known in the art. In such competitive binding assays the Zmax1 or
15 HBM protein or fragment is typically labeled. Free Zmax1 or HBM protein or fragment is separated from that present in a protein:protein complex, and the amount of free (i.e., uncomplexed) label is a measure of the binding of the agent being tested to Zmax1 or HBM or its interference with Zmax1 or HBM: ligand binding, respectively.

Another technique for drug screening provides high throughput screening for
20 compounds having suitable binding affinity to the Zmax1 or HBM proteins and is described in detail in WO 84/03564. Briefly stated, large numbers of different small peptide test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. The peptide test compounds are reacted with Zmax1 or HBM proteins and washed. Bound Zmax1 or HBM protein is then detected by methods well known in the art. Purified Zmax1

or HBM can be coated directly onto plates for use in the aforementioned drug screening techniques. However, non-neutralizing antibodies to the protein can be used to capture antibodies to immobilize the Zmax1 or HBM protein on the solid phase.

This invention also contemplates the use of competitive drug screening assays in which neutralizing antibodies capable of specifically binding the Zmax1 or HBM protein compete with a test compound for binding to the Zmax1 or HBM protein or fragments thereof. In this manner, the antibodies can be used to detect the presence of any peptide that shares one or more antigenic determinants of the Zmax1 or HBM protein.

A further technique for drug screening involves the use of host eukaryotic cell lines or cells (such as described above) that have a nonfunctional Zmax1 or HBM gene. These host cell lines or cells are defective at the Zmax1 or HBM protein level. The host cell lines or cells are grown in the presence of drug compound. The rate of growth of the host cells or impact on lipid metabolism is measured to determine if the compound is capable of regulating the growth or lipid metabolism of Zmax1 or HBM defective cells.

The goal of rational drug design is to produce structural analogs of biologically active proteins of interest or of small molecules with which they interact (e.g., agonists, antagonists, inhibitors) in order to fashion drugs which are, for example, more active or stable forms of the protein, or which, e.g., enhance or interfere with the function of a protein *in vivo*. See, e.g., Hodgson, *Bio/Technology*, 9:19-21 (1991). In one approach, one first determines the three-dimensional structure of a protein of interest (e.g., Zmax1 or HBM protein) or, for example, of the Zmax1- or HBM-receptor or ligand complex, by X-ray crystallography, by computer modeling or most typically, by a combination of approaches. Less often, useful information regarding the structure of a protein may be gained by modeling based on the structure of homologous proteins. An example of rational drug design is the development of

HIV protease inhibitors (Erickson *et al.*, *Science*, 249:527-533 (1990)). In addition, peptides (e.g., Zmax1 or HBM protein) are analyzed by an alanine scan (Wells, *Methods in Enzymol.*, 202:390-411 (1991)). In this technique, an amino acid residue is replaced by Ala, and its effect on the peptide's activity is determined. Each of the amino acid residues of the peptide is analyzed in this manner to determine the important regions of the peptide.

It is also possible to isolate a target-specific antibody, selected by a functional assay, and then to solve its crystal structure. In principle, this approach yields a pharmacore upon which subsequent drug design can be based. It is possible to bypass protein crystallography altogether by generating anti-idiotypic antibodies (anti-ids) to a functional, pharmacologically active antibody. As a mirror image of a mirror image, the binding site of the anti-ids would be expected to be an analog of the original receptor. The anti-id could then be used to identify and isolate peptides from banks of chemically or biologically produced banks of peptides. Selected peptides would then act as the pharmacore.

Thus, one may design drugs which have, e.g., improved Zmax1 or HBM protein activity or stability or which act as inhibitors, agonists, antagonists, etc. of Zmax1 or HBM protein activity. By virtue of the availability of cloned Zmax1 or HBM sequences, sufficient amounts of the Zmax1 or HBM protein may be made available to perform such analytical studies as X-ray crystallography. In addition, the knowledge of the Zmax1 or HBM protein sequence provided herein will guide those employing computer modeling techniques in place of, or in addition to x-ray crystallography.

XIX. Methods of Use: Avian and Mammalian Animal Husbandry

The Zmax1 DNA and Zmax1 protein and/or the HBM DNA and HBM protein can be used for vertebrate and preferably human therapeutic agents and for avian and mammalian

veterinary agents, including for livestock breeding. Animals contemplated as subjects include livestock (e.g., cattle, pigs, sheep, goats, horses, buffalo, etc.), primates, canines, felines, rodents, birds, as well as reptiles, fish, and amphibians. Birds, including, for example, chickens, roosters, hens, turkeys, ostriches, ducks, pheasants and quails, can benefit from the identification of the gene and pathway for high bone mass. In many examples cited in literature (for example, McCoy *et al.*, *Res. Vet. Sci.*, 60(2):185-186 (1996)), weakened bones due to husbandry conditions cause cage layer fatigue, osteoporosis and high mortality rates. Additional therapeutic agents to treat osteoporosis or other bone disorders in birds can have considerable beneficial effects on avian welfare and the economic conditions of the livestock industry, including, for example, meat and egg production.

XX. Methods of use: Diagnostic assays using Zmax1-specific oligonucleotides for detection of genetic alterations affecting bone development and lipid regulation.

In cases where an alteration or disease of bone development or lipid metabolism is suspected to involve an alteration of the *Zmax1* gene or the *HBM* gene, specific oligonucleotides may be constructed and used to assess the level of *Zmax1* mRNA or *HBM* mRNA, respectively, in bone tissue or in another tissue that affects bone development.

For example, to test whether a person has the *HBM* gene, which affects bone density and lipid regulation, polymerase chain reaction can be used. Two oligonucleotides are synthesized by standard methods or are obtained from a commercial supplier of custom-made oligonucleotides. The length and base composition are determined by standard criteria using the Oligo 4.0 primer Picking program (Wojchich Rychlik, 1992). One of the oligonucleotides is designed so that it will hybridize only to *HBM* DNA under the PCR conditions used. The other oligonucleotide is designed to hybridize a segment of *Zmax1*

genomic DNA, such that amplification of DNA using these oligonucleotide primers produces a conveniently identified DNA fragment. For example, the pair of primers CCAAGTTCTGAGAAGTCC (SEQ ID NO:32) and AATACCTGAAACCA TACCTG (SEQ ID NO:33) will amplify a 530 base pair DNA fragment from a DNA sample when the following conditions are used: step 1 at 95°C for 120 seconds; step 2 at 95°C for 30 seconds; step 3 at 58°C for 30 seconds; step 4 at 72°C for 120 seconds; where steps 2-4 are repeated 35 times. Tissue samples may be obtained from hair follicles, whole blood, or the buccal cavity.

The fragment generated by the above procedure is sequenced by standard techniques. Individuals heterozygous for the *HBM* gene will show an equal amount of G and T at the second position in the codon for glycine 171. Normal or homozygous wild-type individuals will show only G at this position.

Other amplification techniques besides PCR may be used as alternatives, such as ligation-mediated PCR or techniques involving Q-beta replicase (Cahill *et al.*, *Clin. Chem.*, 37(9):1482-5 (1991)). For example, the oligonucleotides AGCTGCTCGTAGCTGTCTCT CCCTGGATCACGGGTACATGTACTGGACAGACTGGGT (SEQ ID NO:34) and TGAGACGCCCCGGATTGAGCGGGCAGGGATAGCTTATTCCCTGTGCCGCATTACG GC (SEQ ID NO:35) can be hybridized to a denatured human DNA sample, treated with a DNA ligase, and then subjected to PCR amplification using the primer oligonucleotides AGCTGCTCGTAG CTGTCTCTCCCTGGA (SEQ ID NO:36) and GCCGTAATGCGGCACAGGGAATAAGCT (SEQ ID NO:37). In the first two oligonucleotides, the outer 27 bases are random sequence corresponding to primer binding sites, and the inner 30 bases correspond to sequences in the *Zmax1* gene. The T at the end of the first oligonucleotide corresponds to the *HBM* gene. The first two oligonucleotides are

ligated only when hybridized to human DNA carrying the *HBM* gene, which results in the formation of an amplifiable 114 bp DNA fragment.

Products of amplification can be detected by agarose gel electrophoresis, quantitative hybridization, or equivalent techniques for nucleic acid detection known to one skilled in the art of molecular biology (Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring, NY (1989)).

Other alterations in the *Zmax1* gene or the *HBM* gene may be diagnosed by the same type of amplification-detection procedures, by using oligonucleotides designed to identify those alterations. These procedures can be used in animals as well as humans to identify alterations in *Zmax1* or *HBM* that affect bone development and/or lipid metabolism or levels.

Expression of *Zmax1* or *HBM* in bone tissue may be accomplished by fusing the cDNA of *Zmax1* or *HBM*, respectively, to a bone-specific promoter in the context of a vector for genetically engineering vertebrate cells. DNA constructs are introduced into cells by packaging the DNA into virus capsids, by the use of cationic liposomes, electroporation, or by calcium phosphate transfection. Transfected cells, preferably osteoblasts, may be studied in culture or may be introduced into bone tissue in animals by direct injection into bone or by intravenous injection of osteoblasts, followed by incorporation into bone tissue (Ko *et al.*, *Cancer Research*, 56(20):4614-9 (1996)). For example, the osteocalcin promoter, which is specifically active in osteoblasts, may be used to direct transcription of the *Zmax1* gene or the *HBM* gene. Any of several vectors and transfection methods may be used, such as retroviral vectors, adenovirus vectors, or vectors that are maintained after transfection using cationic liposomes, or other methods and vectors described herein.

Similarly *Zmax1*, or *HBM* can be expressed in liver tissue or in other lipid-metabolism or lipid-regulating cells, such as lipid laden foam cells or lesion cells. This can

be accomplished by fusing the cDNA of Zmax1 or HBM respectively to, for example, a liver specific promoter or other suitable promoter in the context of a vector for genetically engineering vertebrate cells. DNA constructs are introduced into cells by packaging the DNA into, for example, virus capsids, by the use of cationic liposomes, electroporation, or calcium phosphate transfection. The transfected cells, preferably liver cells, may be studied in culture or can be introduced into animals by direct injection into the liver or other cell involved in lipid regulation or metabolism. The vectors and transfection methods to be used are similar to those described herein.

Alteration of the level of functional Zmax1 protein or HBM protein affects the level of bone mineralization and lipid levels. By manipulating levels of functional Zmax1 protein or HBM protein, it is possible to affect bone development and to increase or decrease levels of bone mineralization as well as lipid levels. For example, it may be useful to increase bone mineralization in patients with osteoporosis. Alternatively, it may be useful to decrease bone mineralization in patients with osteopetrosis or Paget's disease. Alteration of Zmax1 levels or HBM levels can also be used as a research tool. Specifically, it is possible to identify proteins, mRNA and other molecules whose level or modification status is altered in response to changes in functional levels of Zmax1 or HBM. The pathology and pathogenesis of bone disorders is known and described, for example, in Rubin and Farber (Eds.), *Pathology*, 2nd Ed., S.B. Lippincott Co., Philadelphia, PA (1994).

Zmax1 or HBM protein levels can be altered to regulate lipid levels in a cell or a subject. The pathology and pathogenesis of atherosclerosis and arteriosclerosis is known and described, for example, in Edwin L. Bierman, "Atherosclerosis and Other Forms of Arteriosclerosis," in *Harrison's Principles of Internal Medicine*, 1106-1116 (13th Ed., 1994). Modulation of lipid levels may be useful to lower certain levels of lipids (e.g., LDL) in

patients with arteriosclerosis and/or atherosclerosis, as well as conditions and diseases affiliated with atherosclerosis and arteriosclerosis, as described by Bierman (1994).

A variety of techniques can be used to alter the levels of functional Zmax1 or HBM. For example, intravenous or intraosseous injection of the extracellular portion of Zmax1 or mutations thereof, or HBM or mutations thereof, will alter the level of Zmax1 activity or HBM activity, respectively, in the body of the treated human, animal or bird. Truncated versions of the Zmax1 protein or HBM protein can also be injected to alter the levels of functional Zmax1 protein or HBM protein, respectively. Certain forms of Zmax1 or HBM enhance the activity of endogenous protein, while other forms are inhibitory.

10 In a preferred embodiment, the HBM protein is used to treat osteoporosis or arteriosclerosis. In a further preferred embodiment, the extracellular portion of the HBM protein is used. This HBM protein may be optionally modified by the addition of a moiety that causes the protein to adhere to the surface of cells. The protein is prepared in a pharmaceutically acceptable solution and is administered by injection or another method that achieves acceptable pharmacokinetics and distribution.

15 In a second embodiment of this method, Zmax1 or HBM levels are increased or decreased by gene therapy techniques. To increase Zmax1 or HBM levels, osteoblasts or another useful cell type are genetically engineered to express high levels of Zmax1 or HBM as described above. Alternatively, to decrease Zmax1 or HBM levels, antisense constructs that specifically reduce the level of translatable Zmax1 or HBM mRNA can be used. In general, a tissue-nonspecific promoter may be used, such as the CMV promoter or another commercially available promoter found in expression vectors (Wu *et al.*, *Toxicol. Appl. Pharmacol.*, 141(1):330-9 (1996)). In a preferred embodiment, a Zmax1 cDNA or its antisense is transcribed by a bone-specific promoter, such as the osteocalcin or another

promoter, to achieve specific expression in bone tissue. In this way, if a Zmax1-expressing DNA construct or HBM-expressing construct is introduced into non-bone tissue, it will not be expressed. Similarly, if a liver-specific promoter is used to express the HBM or Zmax1 proteins in liver or other cell involved in lipid regulation or metabolism, the DNA construct
5 with, for example, a liver-specific promoter will not be expressed in other non-liver tissues.

In a third embodiment of this method, antibodies against Zmax1 or HBM are used to inhibit its function. Such antibodies are identified herein.

In a fourth embodiment of this method, drugs that inhibit Zmax1 function or HBM function are used. Such drugs are described herein and optimized according to techniques of
10 medicinal chemistry well known to one skilled in the art of pharmaceutical development.

Zmax1 and HBM interact with several proteins, such as ApoE. Molecules that inhibit the interaction between Zmax1 or HBM and ApoE or another binding partner are expected to alter bone development and mineralization. Such inhibitors may be useful as drugs in the treatment of osteoporosis, osteopetrosis, or other diseases of bone mineralization. Such
15 inhibitors may be low molecular weight compounds, proteins or other types of molecules. See, Kim *et al.*, *J. Biochem.* (Tokyo), 124(6):1072-1076 (1998).

Inhibitors of the interaction between Zmax1 or HBM and interacting proteins may be isolated by standard drug-screening techniques. For example, Zmax1 protein, (or a fragment thereof) or HBM protein (or a fragment thereof) can be immobilized on a solid support such
20 as the base of microtiter well. A second protein or protein fragment, such as ApoE is derivatized to aid in detection, for example with fluorescein. Iodine, or biotin, then added to the Zmax1 or HBM in the presence of candidate compounds that may specifically inhibit this protein-protein domain of Zmax1 or HBM, respectively, and thus avoid problems associated

with its transmembrane segment. Drug screens of this type are well known to one skilled in the art of pharmaceutical development.

Because Zmax1 and HBM are involved in bone development and lipid regulation, proteins that bind to Zmax1 and HBM are also expected to be involved in bone development and lipid regulation. Such binding proteins can be identified by standard methods, such as co-immunoprecipitation, co-fractionation, or the two-hybrid screen (Ausubel *et al.*, *Current Protocols in Molecular Biology*, John Wiley & Sons (1997)). For example, to identify Zmax1-interacting proteins or HBM-interacting proteins using the two-hybrid system, the extracellular domain of Zmax1 or HBM is fused to LexA and expressed for the yeast vector pEG202 (the "bait") and expressed in the yeast strain EGY48. The yeast strain is transformed with a "prey" library in the appropriate vector, which encodes a galactose-inducible transcription-activation sequence fused to candidate interacting proteins. The techniques for initially selecting and subsequently verifying interacting proteins by this method are well known to one skilled in the art of molecular biology (Ausubel *et al.*, *Current Protocols in Molecular Biology*, John Wiley & Sons (1997)).

In a preferred embodiment, proteins that interact with HBM, but not Zmax1, are identified using a variation of the above procedure (Xu *et al.*, *Proc. Natl. Acad. Sci. USA*, 94(23):12473-8 (Nov. 1997)). This variation of the two-hybrid system uses two baits, and Zmax1 and HBM are each fused to LexA and TetR, respectively. Alternatively, proteins that interact with the HBM but not Zmax1 are also isolated. These procedures are well known to one skilled in the art of molecular biology, and are a simple variation of standard two-hybrid procedures.

As an alternative method of isolating Zmax1 or HBM interacting proteins, a biochemical approach is used. The Zmax1 protein or a fragment thereof, such as the

extracellular domain, or the HBM protein or a fragment thereof, such as the extracellular domain, is chemically coupled to Sepharose beads. The Zmax1- or HBM-coupled beads are poured into a column. An extract of proteins, such as serum proteins, proteins in the supernatant of a bone biopsy, or intracellular proteins from gently lysed TE85 osteoblastic cells, is added to the column. Non-specifically bound proteins are eluted, the column is washed several times with a low-salt buffer, and then tightly binding proteins are eluted with a high-salt buffer. These are candidate proteins that bind to Zmax1 or HBM, and can be tested for specific binding by standard tests and control experiments. Sepharose beads used for coupling proteins and the methods for performing the coupling are commercially available (Sigma), and the procedures described here are well known to one skilled in the art of protein biochemistry.

As a variation of the above procedure, proteins that are eluted by high salt from the Zmax1- or HBM-Sepharose column are then added to an HBM-Zmax1-sepharose column. Proteins that flow through without sticking are proteins that bind to Zmax1 but not to HBM. Alternatively, proteins that bind to the HBM protein and not to the Zmax1 protein can be isolated by reversing the order in which the columns are used. Similar columns can be prepared for use in assessing lipid regulation in liver and other tissues and cells involved in lipid regulation and or metabolism.

XXI. Method of Use: Transformation-Associated Recombination (TAR) Cloning

Essential for the identification of novel allelic variants of Zmax1 is the ability to examine the sequence of both copies of the gene in an individual. To accomplish this, two "hooks," or regions of significant similarity, are identified within the genomic sequence such that they flank the portion of DNA that is to be cloned. Most preferably, the first of these

hooks is derived from sequences 5' to the first exon of interest and the second is derived from sequences 3' to the last exon of interest. These two "hooks" are cloned into a bacterial/yeast shuttle vector such as that described by Larionov *et al.*, *Proc. Natl. Acad. Sci. USA*, 94:7384-7387 (1997). Other similar vector systems may also be used. To recover the entire genomic

5 copy of the *Zmax1* gene, the plasmid containing the two "hooks" is linearized with a restriction endonuclease or is produced by another method such as PCR. This linear DNA fragment is introduced into yeast cells along with human genomic DNA. Typically, the yeast *Saccharomyces cerevisiae* is used as a host cell, although chicken host cells can be used as well (Larionov *et al.*, *Genet. Eng. (NY)*, 21:37-55 (1999). During and after the process of

10 transformation, the endogenous host cell converts the linear plasmid to a circle by a recombination event whereby the region of the human genomic DNA homologous to the "hooks" is inserted into the plasmid. This plasmid can be recovered and analyzed by methods well known to one skilled in the art. Obviously, the specificity for this reaction requires the host cell machinery to recognize sequences similar to the "hooks" present in the linear

15 fragment. However, 100% sequence identity is not required, as shown by Kouprina *et al.*, *Genomics*, 53(1):21-28 (October 1998), where the author describes using degenerate repeated sequences common in the human genome to recover fragments of human DNA from a rodent/human hybrid cell line.

In another example, only one "hook" is required, as described by Larionov *et al.*,

20 *Proc. Natl. Acad. Sci. USA*, 95(8):4469-74 (April 1998). For this type of experiment, termed "radial TAR cloning," the other region of sequence similarity to drive the recombination is derived from a repeated sequence from the genome. In this way, regions of DNA adjacent to the *Zmax1* gene coding region can be recovered and examined for alterations that may affect function.

XXII. Methods of Use: Genomic Screening

The use of polymorphic genetic markers linked to the *HBM* gene or to *Zmax1* is very useful in predicting susceptibility to osteoporosis or other bone diseases. Polymorphic genetic markers linked to the *HBM* gene or the *Zmax1* gene also can be used to predict susceptibility to arteriosclerosis or atherosclerosis and conditions related thereto. Koller *et al.*, *Amer. J. Bone Min. Res.*, 13:1903-1908 (1998) have demonstrated that the use of polymorphic genetic markers is useful for linkage analysis. Similarly, the identification of polymorphic genetic markers within the *HBM* gene will allow the identification of specific allelic variants that are in linkage disequilibrium with other genetic lesions that affect bone development. Using the DNA sequence from the BACs, a dinucleotide CAn repeat was identified and two unique PCR primers that will amplify the genomic DNA containing this repeat were designed, as shown below:

B200E21C16_L: GAGAGGCTATATCCCTGGGC (SEQ ID NO:38)

B200E21C16_R: ACAGCACGTGTTTAAAGGGG (SEQ ID NO:39)

and used in the genetic mapping study.

This method has been used successfully by others skilled in the art (e.g., Sheffield *et al.*, *Genet.*, 4:1837-1844 (1995); LeBlanc-Straceski *et al.*, *Genomics*, 19:341-9 (1994); Chen *et al.*, *Genomics*, 25:1-8 (1995)). Use of these reagents with populations or individuals will predict their risk for osteoporosis. Similarly, single nucleotide polymorphisms (SNPs), such as those shown in Table 4 above, can be used as well to predict risk for developing bone diseases or resistance to osteoporosis in the case of the *HBM* gene. It is also contemplated that single nucleotide polymorphisms (SNPs) such as those described above, may be used to predict the risk in a subject for developing arteriosclerosis and atherosclerosis and related conditions.

XXIII. Methods of Use: Modulators of Tissue Calcification

The calcification of tissues in the human body is well documented. Towler *et al.*, *J. Biol. Chem.*, 273:30427-34 (1998) demonstrated that several proteins known to regulate calcification of the developing skull in a model system are expressed in calcified aorta. The expression of Msx2, a gene transcribed in osteoprogenitor cells, in calcified vascular tissue indicates that genes which are important in bone development are involved in calcification of other tissues. Treatment with HBM protein, agonists or antagonists is likely to ameliorate calcification (such as the vasculature, dentin and bone of the skull visera) due to its demonstrated effect on bone mineral density. In experimental systems where tissue calcification is demonstrated, the over-expression or repression of *Zmax1* activity permits the identification of molecules that are directly regulated by the *Zmax1* gene. These genes are potential targets for therapeutics aimed at modulating tissue calcification. For example, an animal, such as the LDLR $-/-$, mouse is fed a high fat diet and is observed to demonstrate expression of markers of tissue calcification, including *Zmax1*. These animals are then treated with antibodies to *Zmax1* or HBM protein, antisense oligonucleotides directed against *Zmax1* or HBM cDNA, or with compounds known to bind the *Zmax1* or HBM protein or its binding partner or ligand. RNA or proteins are extracted from the vascular tissue and the relative expression levels of the genes expressed in the tissue are determined by methods well known in the art. Genes that are regulated in the tissue are potential therapeutic targets for pharmaceutical development as modulators of tissue calcification.

The nucleic acids, proteins, peptides, amino acids, small molecules or other pharmaceutically useful compounds of the present invention that are to be given to an individual may be administered in the form of a composition with a pharmaceutically acceptable carrier, excipient or diluent, which are well known in the art. The individual may

be a mammal or a bird, preferably a human, a rat, a mouse or bird. Such compositions may be administered to an individual in a pharmaceutically effective amount. The amount administered will vary depending on the condition being treated and the patient being treated. The compositions may be administered alone or in combination with other treatments.

5 XXIV. Pharmaceutical Compositions

The invention also contemplates pharmaceutical compositions comprising a lipid mediating agent which modulates HBM and/or Zmax1 activity in combination with a lipoprotein modulating agent (e.g., blofibrate, gemfibrozil, nicotinic acid, cholestyramine, cholestipol, lovastatin, simvastatin, pravastatin, probucol, premarin or estradiol.). Liprotein
10 modulating agents can include compounds or compositions which modulate (e.g., up-regulate or down-regulate) LDL, VLDL, HDL or IDL levels.

The lipid mediating agent, which modulates HBM and/or Zmax1 activity, can include proteins, monoclonal antibodies or fragments thereof, chemicals, and mimetics. One contemplated pharmaceutical composition can comprise the monoclonal antibody and a
15 pharmaceutically acceptable carrier. For the purposes of the present invention, a "pharmaceutically acceptable carrier" can be any of the standard carriers well known in the art. For example, suitable carriers can include phosphate buffered saline solutions, emulsions such as oil/water emulsions, and various types of wetting agents. Other carriers can also include sterile solutions, tablets, coated tablets, and capsules. Typically, such carriers can
20 also contain excipients such as starch, milk, sugar, types of clay, gelatin, stearic acid, or salts thereof, magnesium or calcium stearate, talc, vegetable fats or oils, gums, glycerols, or other known excipients. Such carriers can also include flavors and color additives, preservatives,

or other ingredients. Compositions comprising such carriers are formulated by well known conventional means. See REMINGTON'S PHARMACEUTICAL SCIENCE (15th ed. 1980).

For diagnostic purposes, the antibodies and recombinant binding proteins can be either labeled or unlabeled. Typically, diagnostic assays entail detecting the formation of a complex through the binding of the monoclonal antibody or recombinant binding protein to a HBM protein or Zmax1 protein. When unlabeled, the antibodies and recombinant binding proteins find use in agglutination assays. In addition, unlabeled antibodies can be used in combination with other labeled antibodies (second antibodies) that are specifically reactive with the monoclonal antibody or recombinant binding protein, such as antibodies specific for immunoglobulin. Alternatively, the monoclonal antibodies and recombinant binding proteins can be directly labeled. A wide variety of labels can be employed, such as radionuclides (e.g., ^{99}Tc , ^{111}In , ^{123}I and ^{131}I), fluorescers, enzymes, enzyme substrates, enzyme cofactors, enzyme inhibitors, ligands (particularly haptens), *etc.* Numerous types of immunoassays are well known in the art.

Commonly, the monoclonal antibodies and recombinant binding proteins of the present invention are used in fluorescent assays, where the subject antibodies or recombinant binding proteins are conjugated to a fluorescent molecule, such as fluorescein isothiocyanate (FITC).

The examples provided below are not meant to limit the invention in any way, but serve to provide preferred embodiments for the invention.

EXAMPLES

The present invention is described by reference to the following Examples, which are offered by way of illustration and are not intended to limit the invention in any manner.

Standard techniques well known in the art or the techniques specifically described below were utilized.

Example 1

The propositus was referred by her physicians to the Creighton Osteoporosis Center
5 for evaluation of what appeared to be unusually dense bones. She was 18 years old and came
to medical attention two years previous because of back pain, which was precipitated by an
auto accident in which the car in which she was riding as a passenger was struck from behind.
Her only injury was soft tissue injury to her lower back that was manifested by pain and
muscle tenderness. There was no evidence of fracture or subluxation on radiographs. The
10 pain lasted for two years, although she was able to attend school full time. By the time she
was seen in the Center, the pain was nearly resolved and she was back to her usual activities
as a high school student. Physical exam revealed a normal healthy young woman standing 66
inches and weighing 128 pounds. Radiographs of the entire skeleton revealed dense looking
bones with thick cortices. All bones of the skeleton were involved. Most importantly, the
15 shapes of all the bones were entirely normal. The spinal BMC was 94.48 grams in L1-4, and
the spinal BMD was 1.667 gm/cm² in L1-4. BMD was 5.62 standard deviations (SD) above
peak skeletal mass for women. These were measured by DXA using a Hologic 2000~. Her
mother was then scanned and a lumbar spinal BMC of 58.05 grams and BMD of 1.500
gm/cm² were found. Her mother's values place her 4.12 SD above peak mass and 4.98 SD
20 above her peers. Her mother was 51 years old, stood 65 inches and weighed 140 pounds.
Her mother was in excellent health with no history of musculoskeletal or other symptoms.
Her father's lumbar BMC was 75.33 grams and his BMD was 1.118 gm/cm². These values

place him 0.25 SD above peak bone mass for males. He was in good health, stood 72 inches tall, and weighed 187 pounds.

These clinical data suggested that the propositus inherited a trait from her mother, which resulted in very high bone mass, but an otherwise normal skeleton, and attention was
5 focused on the maternal kindred. In U.S. Patent No. 5,691,153, twenty- two of these members had measurement of bone mass by DXA. In one case, the maternal grandfather of the propositus, was deceased, however, medical records, antemortem skeletal radiographs and a gall bladder specimen embedded in paraffin for DNA genotyping were obtained. His radiographs showed obvious extreme density of all of the bones available for examination
10 including the femur and the spine, and he was included among the affected members. In this invention, the pedigree has been expanded to include 37 informative individuals. These additions are a significant improvement over the original kinship (Johnson *et al.*, *Am. J. Hum. Genet.*, 60:1326-1332 (1997)) because, among the fourteen individuals added since the original study, two individuals harbor key crossovers. X-linkage is ruled out by the presence
15 of male-to-male transmission from individual 12 to 14 and 15.

Example 2

The present invention describes DNA sequences derived from two BAC clones from the *HBM* gene region, as evident in Table 6 below, which is an assembly of these clones. Clone b200e21-h (ATCC No. 980812; SEQ ID NOS: 10-11) was deposited at the American
20 Type Culture Collection (ATCC), 10801 University Blvd., Manassas, VA 20110-2209 U.S.A., on December 30, 1997. Clone b527d12-h (ATCC No. 980720; SEQ ID NOS: 5-9) was deposited at the American Type Culture Collection (ATCC), 10801 University Blvd., Manassas, VA 20110-2209 U.S.A., on October 2, 1998. These sequences are unique reagents

that can be used by one skilled in the art to identify DNA probes for the *Zmax1* gene, PCR primers to amplify the gene, nucleotide polymorphisms in the *Zmax1* gene, or regulatory elements of the *Zmax1* gene.

TABLE 6

Contig	ATCC No.	SEQ ID NO.	Length
b527d12-h_contig302G	980720	5	3096
b527d12-h_contig306G	980720	6	26928
b527d12-h_contig307G	980720	7	29430
b527d12-h_contig308G	980720	8	33769
b527d12-h_contig309G	980720	9	72049
b200e21-h_contig1	980812	10	8705
b200e21-h_contig4	980812	11	66933

The disclosure of each of the patents, patent applications and publications cited in the specification is hereby incorporated by reference herein in its entirety.

Although the invention has been set forth in detail, one skilled in the art will recognize that numerous changes and modifications can be made, and that such changes and modifications may be made without departing from the spirit and scope of the invention.

Example 3

Since *Zmax1* has similarity to the LDL receptor family of genes, it may be involved in lipid metabolism. However, others have reported that lipid profile variables did not show significant association with bone mass and could not be used as indicators for bone mineral density (Zabaglia *et al.*, "An exploratory study of association between lipid profile and bone mineral density in menopausal women in a Campinas reference hospital," *Cad. Saude Publica* 14: 779-86 (1998)). *Zmax1* may be normally involved in regulating bone density by

depositing calcium during bone remodeling. The HBM mutation may result in increased deposition thus conferring denser bone structure. Interestingly, atherosclerotic plaques contain calcified material and express a variety of genes involved in bone differentiation.

To test whether the HBM gene was involved in lipid regulation, biochemical tests were performed to measure serum level of various lipid containing molecules or precursors in affected and unaffected HBM family members to test whether the HBM mutation in the *Zmax1* gene effects lipid metabolism. Table 7 shows the results of testing eight HBM individuals and seven unaffected individuals. Wilcoxon rank-sum tests (non-parametric equivalent of a T-test) were performed to assess whether levels of biochemical markers from affected HBM individuals deviated from unaffected individuals. The data obtained were analyzed separately by gender, as well as by combining values from males and females, when appropriate.

Standard diagnostic protocols were used to determine the concentration (mg/dL) with triglycerides, cholesterol, high density lipoprotein (HDL), low density lipoprotein (LDL), very low density lipoprotein (VLDL), apolipoprotein A-1 (APO A-1), apolipoprotein B (APO B), and lipoprotein a (LIPoa). For such procedures, see for example, F. W. HEMMING, LIPID ANALYSIS (Bios Scientific Pub. 1996) and J. M. ORDOVAS, LIPOPROTEIN PROTOCOLS (Humana Press Inc., 1997). The genotype for apolipoprotein E (APO E) was also reported. There are three common alleles (*e.g.*, E2, E3 and E4). The affected and unaffected HBM family members are heterozygous or homozygous for the alleles.

The results obtained were statistically significant: (1) Triglyceride levels were generally lower in affected individuals than in unaffected individuals, and (2) very low density lipoprotein (VLDL) levels were generally lower in affected individuals than in unaffected individuals. Additionally, the following comparisons approached statistical significance ($p=0.06$): (1) high density lipoprotein (HDL) levels were higher in affected males than in unaffected males, and (2) the ratio of low density lipoprotein (LDL) to high density lipoprotein (HDL) was generally higher in affected males than in unaffected males.

In Table 7, "ARUP" is ARUP Laboratories, 500 Chipeta Way, Salt Lake City, UT 84108 where one of the studies was performed. "SJH" refers to the second center which

performed these studies, Creighton Medical Laboratories, 28th & Burt, Dental-Rm 306, Omaha, NE 68178. APO-A1, APO-B and LIPO-a are reported in mg.dL. Total serum levels also are in mg/dL.

All cited patents and publications referred to in this application are herein
5 incorporated by reference in their entirety.

Lipid Studies in HBM Kindred

ID Numbers	Z-score	Gender	TRIGLYC, mg/dl		CHOLESTEROL, mg/dl		HDL, mg/dl		VLDL, mg/dl		LDL, mg/dl		TC-HDL		LDL-HDL		APO A-1	APO B	LIPO a	APO E
			ARUP	SJH	ARUP	SJH	ARUP	SJH	ARUP	SJH	ARUP	SJH	ARUP	SJH	ARUP	SJH				
HBM-ORI-9018	2.36	M	105	105	111	117	48	47	21	21	42	49	2.31	2.50	0.88	1.00	133	57	6	2/4
HBM-ORI-1044	6.05	F	125	129	208	207	67	68	25	26	126	123	3.65	3.60	2.21	2.10	179	125	15	3/4
HBM-ORI-1041	6.21	F	88	88	175	175	86	82	18	18	71	75	2.03	2.10	0.83	0.80	220	74	8	2/3
HBM-ORI-1031	3.42	M	69	76	159	167	49	47	14	15	96	105	3.24	3.60	1.96	2.20	142	96	4	2/3
HBM-ORI-0115	6.47	M	113	113	229	219	70	69	23	23	136	127	3.27	3.20	1.94	1.80	192	122	10	3/3
HBM-ORI-0114	3.31	M	373	370	240	241	64	55	75	74	111	112	4.44	4.40	2.06	2.00	167	134	4	2/3
HBM-ORI-0113	3.70	F	107	104	134	139	72	68	21	21	41	50	1.86	2.00	0.57	0.70	179	55	5	2/3
HBM-ORI-0001	5.92	F	109	113	152	155	44	44	22	23	86	86	3.45	3.50	1.95	2.00	150	95	9	3/3
HBM-ORI-1042	1.85	M	125	120	163	163	44	40	25	24	94	89	3.70	4.10	2.14	2.50	130	106	17	3/4
HBM-ORI-1033	0.99	F	258	282	212	225	43	46	52	56	117	123	4.83	4.80	2.72	2.70	162	126	14	2/3
HBM-ORI-1030		M	363	361	231	244	40	38	73	72	118	134	5.78	6.40	2.95	3.50	138	139	8	2/3
HBM-ORI-1028	1.05	F	186	217	168	182	51	55	37	43	80	84	3.29	3.30	1.57	1.50	189	112	6	3/3
HBM-ORI-1026	-0.21	F	160	175	244	256	38	38	32	35	174	183	6.42	6.70	4.58	4.80	129	161	14	3/3
HBM-ORI-1025	-0.57	F	218	240	197	207	57	58	44	46	96	101	3.46	3.60	1.68	1.70	189	111	8	2/3
HBM-ORI-0148	0.45	M	188	189	188	183	42	44	38	38	108	111	4.48	4.40	2.57	2.50	140	120	2	3/4
		Mean A	135	137	176	176	60	69	27	28	89	91	3.03	3.11	1.55	1.59	170	95	7.63	
		St Dev A	97	95	46	42	15	13	20	19	36	31	0.89	0.84	0.67	0.61	29	31	3.74	
		Mean U	214	226	200	210	45	46	43	45	112	119	4.58	4.77	2.60	2.74	154	125	9.86	
		St Dev U	78	78	31	34	7	8	16	16	30	33	1.20	1.32	1.01	1.12	26	19	5.30	
		P	0.11	0.07	0.25	0.13	0.03	0.04	0.12	0.08	0.19	0.11	0.01	0.01	0.01	0.03	0.27	0.04	0.36	

Cognition
Analysis

GTC Analysis	P-value	Combined	0.02		0.22		0.02		0.32		0.34		0.20		0.34		0.68		1.00	
	P-value	Female	0.03		0.15		0.03		0.34		1.00		0.11		0.08		0.23		0.85	
	P-value	Male	0.40		1.00		0.40		1.00											

CLAIMS

What is claimed is:

1. A method of identifying a molecule involved in lipid regulation comprising identifying a molecule that binds to, or that inhibits binding of a molecule to, HBM or *Zmax1*.
2. The method of claim 1, wherein said molecule is a protein.
3. The method of claim 2, further comprising producing an antibody to the protein.
4. A method for identifying a protein involved in lipid regulation comprising identifying a protein that has an expression level that is different in a first host comprising the *Zmax1* gene when compared to a second host comprising the *HBM* gene.
5. The method of claim 4, wherein the host is an animal.
6. A method for identification of a candidate molecule involved in lipid regulation comprising:
 - (A) identifying a molecule that binds to, or that inhibits binding of a molecule to, the nucleic acid sequence of SEQ ID NO: 1 or a *Zmax1* nucleic acid comprising a polymorphism of Table 4;
 - (B) identifying a molecule that binds to, or that inhibits binding of a molecule to, the nucleic acid sequence of SEQ ID NO: 2; and
 - (C) comparing the extent of binding, or the extent of inhibition of binding, of the molecule to each nucleic acid sequence, wherein the molecule that binds, or inhibits binding, more or less to the nucleic acid sequence of SEQ ID NO: 2 or the nucleic acid sequence of

SEQ ID NO: 1 or a Zmax1 nucleic acid comprising a polymorphism of Table 4 is the candidate molecule.

7. The method of claim 6, wherein the candidate molecule is a protein, an mRNA or an antisense nucleic acid.

8. A method for testing a substance as a therapeutic agent for modulating lipid levels comprising administering a nucleic acid comprising SEQ ID NO: 2 or a nucleic acid sequence with an HBM polymorphism to a subject, and assessing whether lipid levels are modulated.

9. The method of claim 8, wherein the subject is an animal and the animal is selected from the group consisting of: livestock, primates, humans, canines, felines, rodents, birds, reptiles, fish, and amphibians.

10. A method for testing a substance as a therapeutic agent for modulating lipid levels comprising administering a protein comprising SEQ ID NO: 4 or a Zmax1 protein comprising a polymorphism of Table 4 to a subject, and assessing whether lipid levels are modulated.

11. A method of pharmaceutical development for treating lipid-mediated disorders comprising identifying a molecule that binds to the amino acid sequence of SEQ ID NO: 4 or to a Zmax1 protein comprising a polymorphism of Table 4.

12. The method of claim 11, wherein the molecule inhibits or enhances the function of the amino acid.

13. A method of pharmaceutical development for treatment of lipid-mediated disorders comprising:

- (A) constructing a first host that contains the *Zmax1* gene or protein;
- (B) constructing a second host that contains the *HBM* gene or protein;
- (C) analyzing a difference between the first host and the second host; and
- (D) identifying a molecule that, when added to the first host, causes the first host to exhibit a characteristic feature of the second host.

14. The method of claim 13, wherein the host is a cell-free extract, a cell or an animal.

15. The method of claim 13, wherein the difference is a surrogate marker.

16. A method of regulating lipid levels in a host comprising administering the amino acid sequence comprising SEQ ID NO: 4 to a somatic cell or to a germ-line cell of a host suffering from a lipid-mediated disorder.

17. The method of claim 16, wherein the host is livestock, primates, humans, canines, felines, rodents, birds, reptiles, fish, or amphibians.

19. A method for treating or preventing a lipid-mediated disorder in an animal comprising transferring a nucleic acid sequence comprising SEQ ID NO: 2 or a *Zmax1* nucleic acid comprising a polymorphism of Table 4 into a somatic cell or a germ-line cell of an animal suffering from a lipid-mediated disorder.

20. The method of claim 19, wherein the animal is livestock, primates, humans, canines, felines, rodents, birds, reptiles, fish, or amphibians.

21. A method of treating or preventing arteriosclerosis or an arteriosclerosis-associated condition comprising administering an amino acid sequence comprising SEQ ID NO: 4 to a patient in need thereof.

22. The method of claim 21, wherein the patient is livestock, primates, humans, canines, felines, rodents, birds, reptiles, fish, or amphibians.

23. The method of claim 21, wherein the amino acid sequence administered to a patient in need thereof comprises the extracellular domain of the amino acid sequence comprising SEQ ID NO: 4.

24. The method of claim 21, wherein the amino acid sequence administered to a patient in need thereof comprises the intracellular domain of the amino acid sequence comprising SEQ ID NO: 4.

25. A method for treating or preventing a lipid-mediated disorders comprising administering a molecule that binds to a nucleic acid sequence comprising SEQ ID NO: 2 or a Zmax1 nucleic acid comprising a polymorphism of Table 4 to a patient in need thereof.

26. The method of claim 25, wherein the patient is livestock, primates, humans, canines, felines, rodents, birds, reptiles, fish, or amphibians.

27. A method for treating or preventing lipid-mediated disorders comprising administering an antibody to a patient in need thereof, wherein the antibody is to the amino acid sequence comprising SEQ ID NO: 4.

28. A method for diagnostic screening for a genetic predisposition to arteriosclerosis or an arteriosclerosis associated condition or a lipid-mediated disorder comprising screening a sample from a patient with a nucleotide sequence derived from the genomic or cDNA nucleic acid sequence of HBM.

29. The method of claim 28, wherein the screening involves performing a haplotype analysis using the nucleic acid sequence comprising SEQ ID NO: 2 and determining whether the subject contains the *Zmax1* gene or lacks an HBM polymorphism.

30. A diagnostic assay for determining a predisposition for a lipid-mediated disorders comprising an antibody to the HBM protein and an antibody to the *Zmax1* protein.

31. A method of expressing the HBM protein in tissue comprising constructing an expression vector comprising a promoter that directs expression in tissue operably linked to SEQ ID NO:2 and the tissue in which the HBM protein is expressed is a lipid regulating cell or a cell involved in lipid metabolism.

32. The method of claim 31, wherein the tissue is liver.

33. The method of claim 31, wherein the promoter that directs expression in tissue is an osteocalcin promoter or an AML-3 promoter.

34. A method of modulating lipid levels in a subject by administering an HBM protein or a *Zmax1* protein comprising a polymorphism of Table 4.

35. The method of claim 34, wherein the HBM protein comprises SEQ ID NO: 4.

36. The method of claim 34, wherein the lipid modulated is selected from the group consisting of: VLDL, LDL, IDL, HDL, LPOa, APO A-1, APO B and APO E.

37. A method of modulating lipid levels in a subject by administering an agent which regulates HBM or *Zmax1* activity.

38. The method of claim 37, wherein the lipid modulated is selected from the group consisting of: VLDL, LDL, IDL, HDL, LIPOa, APO A-1, APO B and APO E.

39. The method of claim 37, wherein the regulation of HBM or Zmax1 activity is modulates gene transcription, protein translation or Zmax1 or HBM protein binding to its cognate target thereby regulating lipid levels.

40. A composition for treating a lipid-mediated condition comprising an agent that modulates lipid levels by regulating Zmax1 or HBM activity and a lipoprotein modulating agent with a pharmaceutically acceptable carrier.

41. The composition of claim 40, wherein the lipoprotein modulating agent is blofibrate, gemfibrozil, nicotinic acid, cholestyramine, cholestipol, lovastatin, simvastatin, pravastatin, probucol, premarin or estradiol.

42. The composition of claim 40, wherein the lipoprotein modulating agent modulates LDL levels.

43. The composition of claim 42, wherein the lipoprotein modulating agent is selected from the group consisting of bile acid binding resins, HMG-CoA reductase inhibitors and estrogens.

44. A method of treating a subject suffering from a lipid-mediated condition comprising the step of administering the composition of claim 40.

45. The method of claim 44, wherein the lipid-mediated condition is atherosclerosis, arteriosclerosis, or a disease associated with atherosclerosis or arteriosclerosis.

46. A combination therapy for treating a subject suffering from a lipid-mediated disease or condition comprising administering to a subject an agent which regulates HBM or Zmax1 and an agent which regulates a lipoprotein.

47. The combination therapy of claim 46, wherein the agent regulating lipoprotein concentrations is blofibrate, gemfibrozil, nicotinic acid, cholestyramine, cholestipol, lovastatin, simvastatin, pravastatin, probucol, premarin or estradiol.

48. The method of claim 46, wherein the lipid-mediated disease is atherosclerosis, arteriosclerosis, an atherosclerosis associated condition or an arteriosclerosis associated condition.

1/29

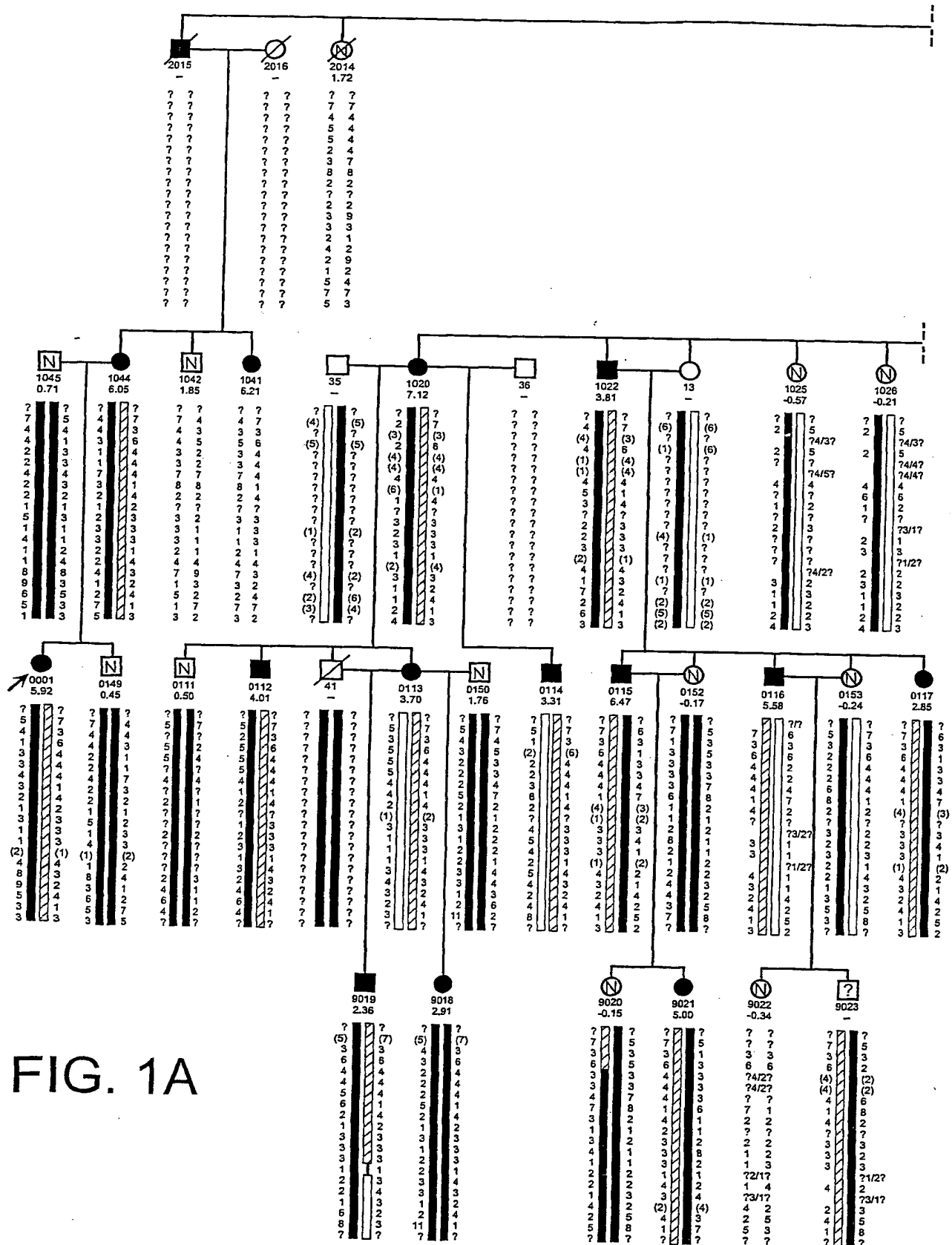


FIG. 1A

2/29

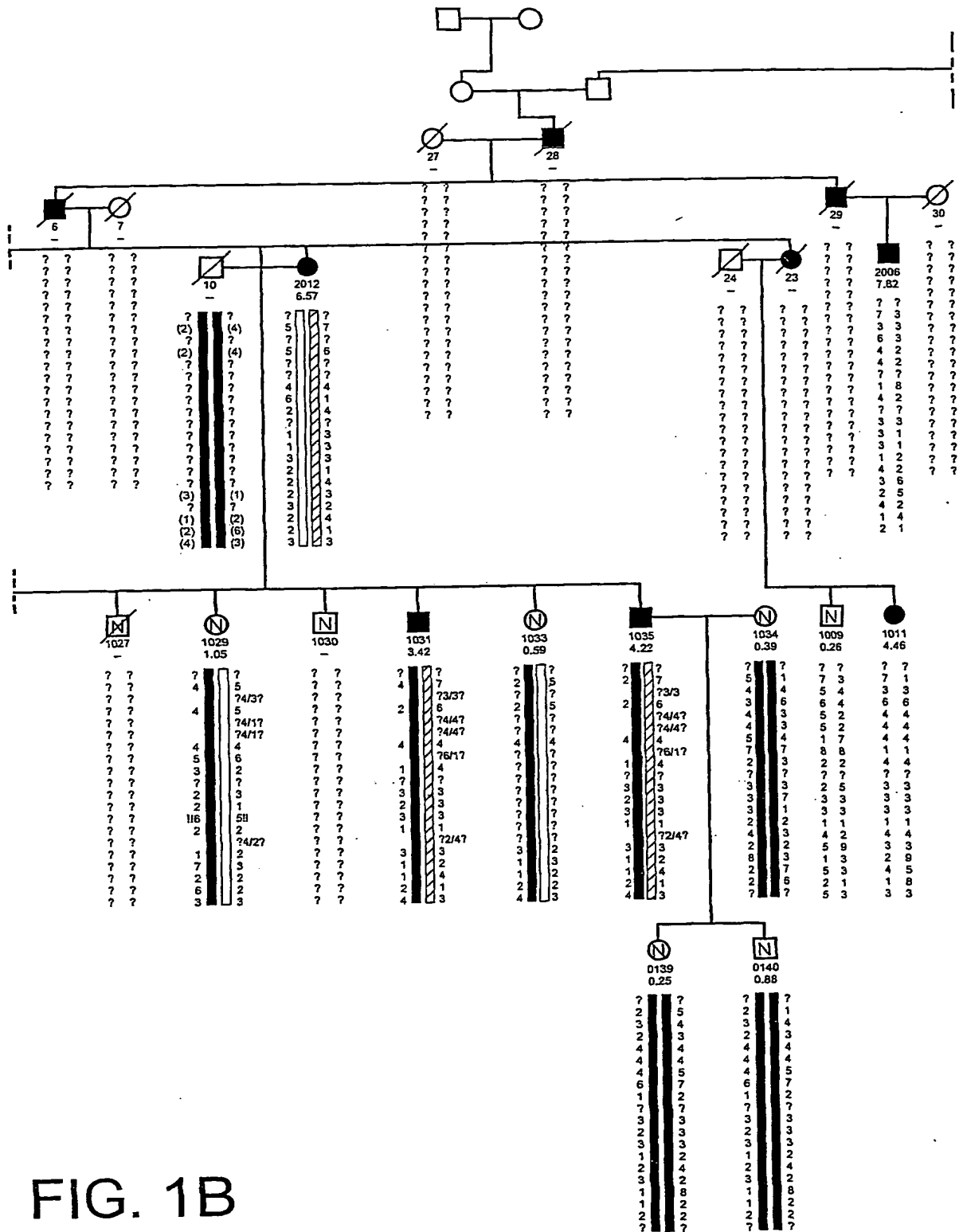
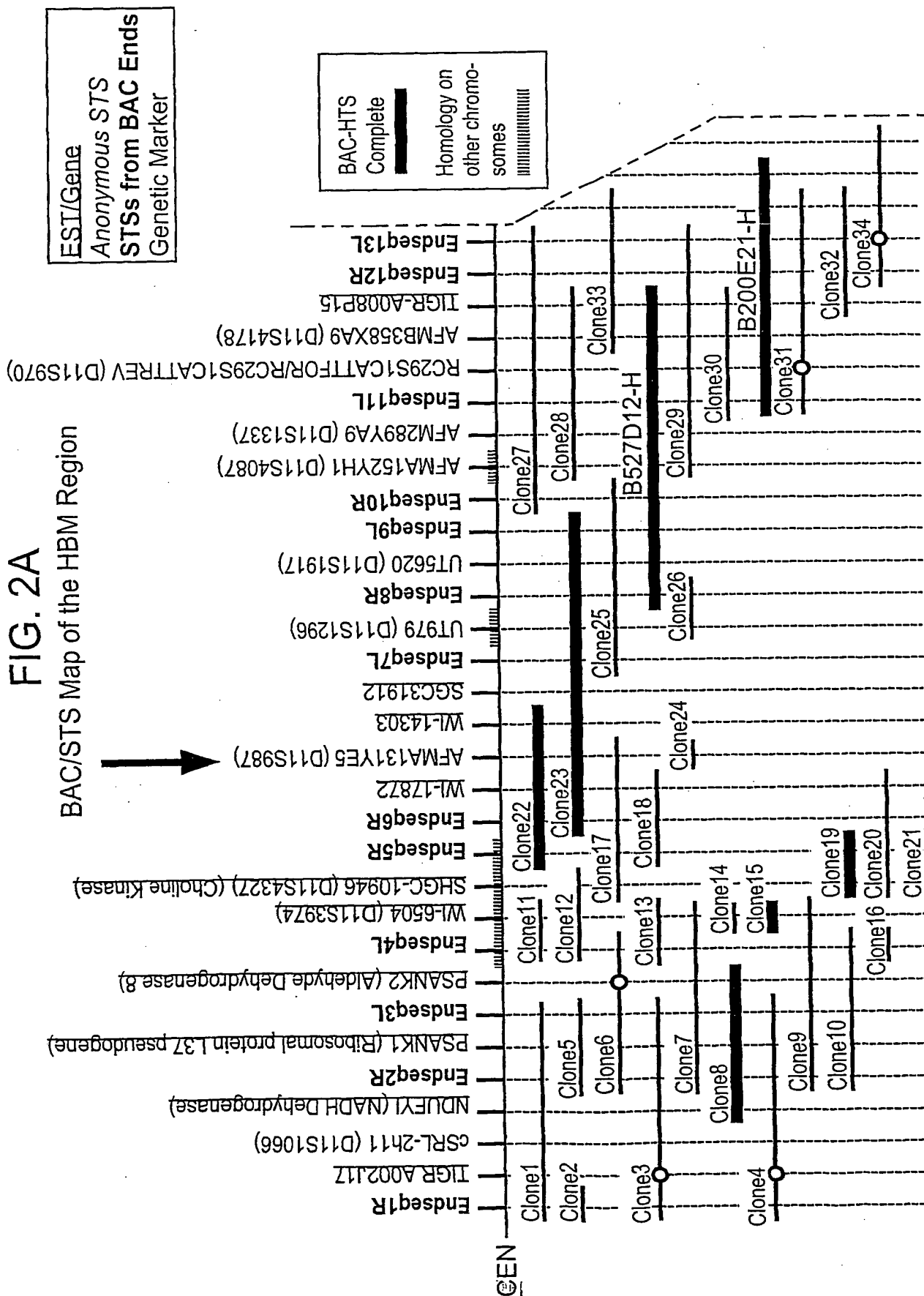


FIG. 1B

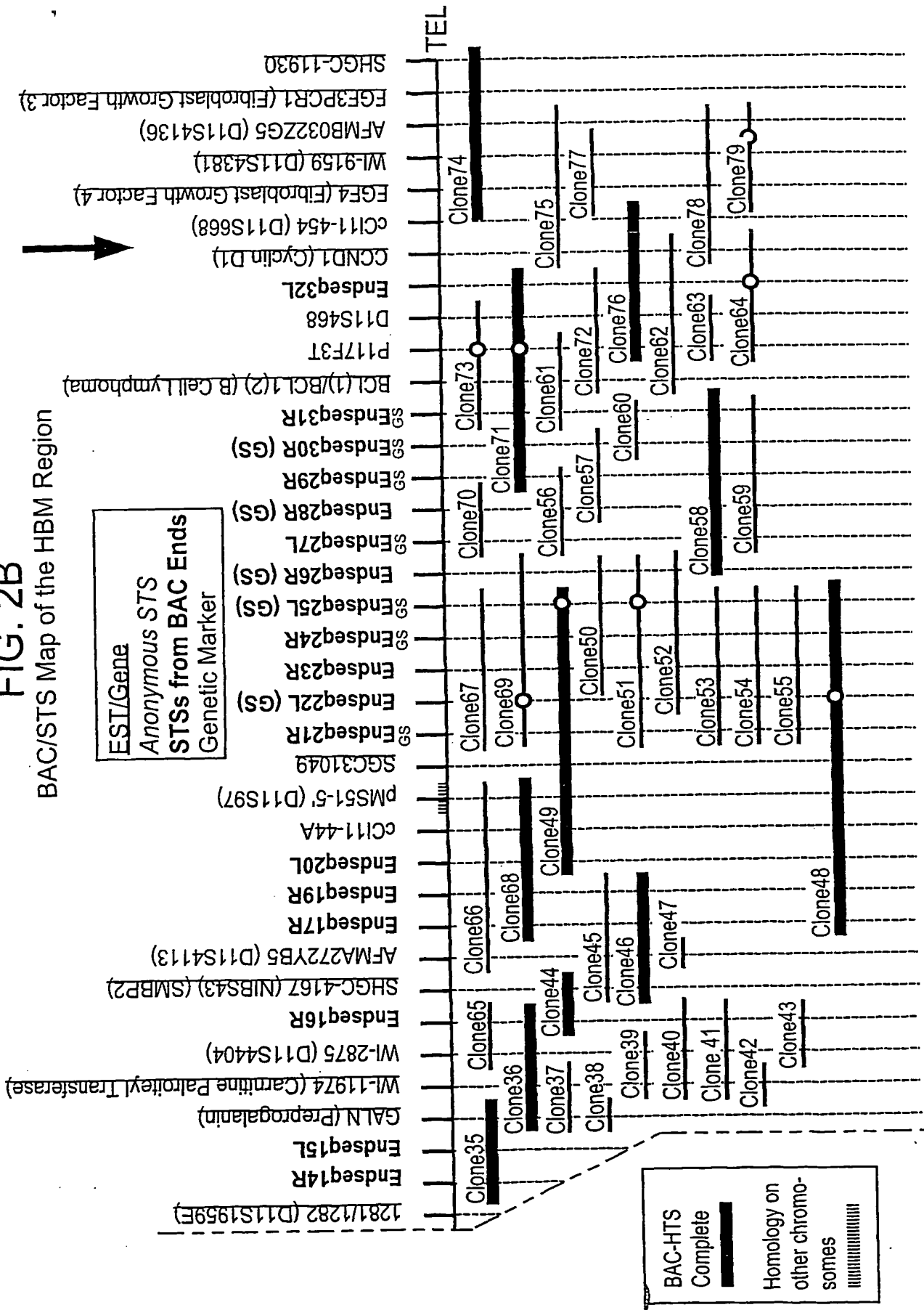
3/29



4/29

FIG. 2B

BAC/STS Map of the HBM Region



5/29

Exon 1

ACTAAAGCGCCGCCGCCGCCATGGAGCCCGAGTGAGCGCGGGCGCGG
GCCCCGTCGGGCCGCCGGACAACATGGAGGCAGCGCCGCCCGGGCCGCC
GTGGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGGCGCTGTGCGGCTGC
CCGGCCCCCGCCGCCGCC

Exon 2 Coordinates: 527d12_Contig308G 30944-30549

gccccacagCCTCGCCGCTCCTGCTATTTGCCAACCGCCGGGACGTACGGCT
GGTGGACGCCGGCGGGAGTCAAGCTGGAGTCCACCATCGTGGTCAGCGG
CCTGGAGGATGCGGCCGCGAGTGGACTTCCAGTTTCCAAGGGAGCCGTG
TACTGGACAGACGTGAGCGAGGAGGCCATCAAGCAGACCTACCTGAACC
AGACGGGGGGCCCGCCGTGCAGAACGTGGTCATCTCCGGCCTGGTCTCTCC
CGACGGCCTCGCCTGCGACTGGGTGGGCAAGAAGCTGTACTGGACGGA
CTCAGAGACCAACCGCATCGAGGTGGCCAACCTCAATGGCACATCCCGG
AAGGTGCTCTTCTGGCAGGACCTTGACCAGCCGAGGGCCATCGCCTTGG
ACCCCGCTCACGGgtaaaccctgctg

... 9408 nt ...

Exon 3 Coordinates: 527d12_Contig308G 21141-20945

ccccgtcacagGTACATGTACTGGACAGACTGGGGGTGAGACGCCCCGGATTGA
GCGGGCAGGGATGGATGGCAGCACCCGGAAGATCATTGTGGACTCGGA
CATTTACTGGCCCAATGGACTGACCATCGACCTGGAGGAGCAGAAGCTC
TACTGGGCTGACGCCAAGCTCAGCTTCATCCACCGTGCCAACCTGGACG
GCTCGTTCCGgtaggtaccac

... 6094 nt ...

Exon 4 Coordinates: 527d12_Contig308G 15047-14850

tccctgactgcagGCAGAAGGTGGTGGAGGGCAGCCTGACGCACCCCTTCGCCC
TGACGCTCTCCGGGGACACTCTGTACTGGACAGACTGGCAGACCCGCTC
CATCCATGCCTGCAACAAGCGCACTGGGGGGAAGAGGAAGGAGATCCTG
AGTGCCCTATACTACCCATGGACATCCAGGTGCTGAGCCAGGAGCGGC
AGCCTTTCTgtgagtgccg

... 1827 nt ...

Exon 5 Coordinates: 527d12_Contig308G 13220-13088

tttctcagTCCACACTCGCTGTGAGGAGGACAATGGCGGCTGCTCCCACCTGT
GCCTGCTGTCCCCAAGCGAGCCTTTCTACACATGCGCCTGCCCCACGGG
TGTGCAGCTGCAGGACAACGGCAGGACGTGTAAGGCAGgtgaggcggtgggacg

FIG. 3A

6/29

... 20923 nt ...

Exon 6 Coordinates: 527d12_Contig309G 7705-8100

ctccacagGAGCCGAGGAGGTGCTGCTGCTGGCCCCGGCGGACGGACCTACG
GAGGATCTCGCTGGACACGCCGGA~~CTT~~CACCGACATCGTGCTGCAGGTG
GACGACATCCGGCACGCCATTGCCATCGACTACGACCCGCTAGAGGGCT
ATGTCTACTGGACAGATGACGAGGTGCGGGGCCATCCGCAGGGCGTACCT
GGACGGGTCTGGGGCGCAGACGCTGGTCAACACCGAGATCAACGACCC
CGATGGCATCGCGGTCTGACTGGGTGGCCCGAAACCTCTACTGGACCGAC
ACGGGCACGGACCGCATCGAGGTGACGCGCCTCAACGGCACCTCCCGCA
AGATCCTGGTGTCGGAGGACCTGGACGAGCCCCGAGCCATCGCACTGCA
CCCCGTGATGGGgtaagacgggc

..... 3211 nt

Exon 7 Coordinates: 527d12_Contig309G 11311-11482

ttcttctccagCCTCATGTACTGGACAGACTGGGGAGAGAACCCTAAAATCGAG
TGTGCCAACTTGATGGGCAGGAGCGGGCGTGTGCTGGTCAATGCCTCCC
TCGGGTGGCCCAACGGCCTGGCCCTGGACCTGCAGGAGGGGGAAGCTCT
ACTGGGGAGACGCCAAGACAGACAAGATCGAGgtgaggctcctgtgg

..... 13445 nt

Exon 8 Coordinates: 527d12_Contig309G 24927-25143

cgtctctgcagGTGATCAATGTTGATGGGACGAAGAGGGCGGACCCTCCTGGAG
GACAAGCTCCCGCACATTTTCGGGTTCACGCTGCTGGGGGACTTCATCT
ACTGGACTGACTGGCAGCGCCGCAGCATCGAGCGGGTGCACAAGGTCAA
GGCCAGCCGGGACGTCATCATTGACCAGCTGCCCGACCTGATGGGGCTC
AAAGCTGTGAATGTGGCCAAGGTCGTCTGgtgagtcgggggggtc

....2826 nt

Exon 9 Coordinates: 527d12_Contig309G 27969-28256

gttcgcttcagGAACCAACCCGTGTGCGGACAGGAACGGGGGGGTGCAGCCACC
TGTGCTTCTTCACACCCACGCAACCCGGTGTGGCTGCCCCATCGGCCT
GGAGCTGCTGAGTGACATGAAGACCTGCATCGTGCCTGAGGCCTTCTTG
GTCTTCACCAGCAGAGCCGCCATCCACAGGATCTCCCTCGAGACCAATA
ACAACGACGTGGCCATCCCGCTCACGGGCGTCAAGGAGGCCTCAGCCCT
GGACTTTGATGTGTCCAACAACCACATCTACTGGACAGACGTCAGCCTG
AAGgtagcgtgggc

.....3102.....

FIG. 3B

7/29

Exon 10 Coordinates: 527d12_Contig309G 31358-31582

cctgctgccagACCATCAGCCGCGCCTTCATGAACGGGAGCTCGGTGGAGCAC
GTGGTGGAGTTTGGCCTTGACTACCCCGAGGGCATGGCCGTTGACTGGA
TGGGCAAGAACCTCTACTGGGCCGACACTGGGACCAACAGAATCGAAGT
GGCGCGGCTGGACGGGCAGTTCCGGCAAGTCCTCGTGTGGAGGGACTT
GGACAACCCGAGGTCGCTGGCCCTGGATCCCACCAAGGGgtaagtgttgccctgc

.....1297 nt.....

Exon 11 Coordinates: 527d12_Contig309G 32879-33064

gtgcctccagCTACATCTACTGGACCGAGTGGGGCGGCAAGCCGAGGATCGT
GCGGGCCTTCATGGACGGGACCAACTGCATGACGCTGGTGGACAAGGTG
GGCCGGGCCAAGCACCCTCACCATTGACTACGCTGACCAGCGCCTCTACT
GGACCGACCTGGACACCAACATGATCGAGTCGTCCAACATGCTGGgtgaggg
ccgggct

.....2069 nt.....

Exon 12 Coordinates: 527d12_Contig309G 35133-35454

gtgttcagcagGTCAGGAGCGGGTCGTGATTGCCGACGATCTCCCGCACCCGT
TCGGTCTGACGCAGTACAGCGATTATATCTACTGGACAGACTGGAATCT
GCACAGCATTGAGCGGGCCGACAAGACTAGCGGCCGGAACCGCACCCCTC
ATCCAGGGGCCACCTGGACTTCGTGATGGACATCCTGGTGTTCCTACTCCT
CCCGCCAGGATGGCCTCAATGACTGTATGCACAACAACGGGGCAGTGTGG
GCAGCTGTGCCTTGCCATCCCCGGCGGCCACCGCTGCGGCTGCGCCTCA
CACTACACCCTGGACCCAGCAGCCGCAACTGCAGCCgtaagtgcctcatgt

.....2006 nt.....

Exon 13 Coordinates: 527d12_Contig309G 37460-37659

gcctcctctaCGCCCACCACCTTCTTGCTGTTTCAGCCAGAAATCTGCCATCAGT
CGGATGATCCCGGACGACCAGCACAGCCCGGATCTCATCCTGCCCTGC
ATGGACTGAGGAACGTCAAAGCCATCGACTATGACCCACTGGACAAGTT
CATCTACTGGGTGGATGGGCGCCAGAACATCAAGCGAGCCAAGGACGAC
GGGACCCAGgcaggtgccctgtgg

.....6965 nt.....

FIG. 3C

8/29

Exon 14 Coordinates: 527d12_Contig309G 44624-44832

ctttgtcttacagCCCTTTGTTTTGACCTCTCTGAGCCAAGGCCAAAACCCAGACA
GGCAGCCCCACGACCTCAGCATCGACATCTACAGCCGGACACTGTTCTG
GACGTGCGAGGCCACCAATACCATCAACGTCCACAGGCTGAGCGGGGAA
GCCATGGGGGTGGTGTCTGCGTGGGGACCGCGACAAGCCCAGGGCCATC
GTCGTCAACGCGGAGCGAGGgtaggaggccaac

.....1404 nt.....

Exon 15 Coordinates: 527d12_Contig309G 46236-46427

ccaccctcccgagGTACCTGTACTTCACCAACATGCAGGACCGGGCAGCCAAGA
TCGAACGCGCAGCCCTGGACGGCACCGAGCGCGAGGTCCTCTTCACCAC
CGGCCTCATCCGCCCTGTGGCCCTGGTGGTGGACAACACACTGGGCAAG
CTGTTCTGGGTGGACGCGGACCTGAAGCGCATTGAGAGCTGTGACCTGT
CAGgtacgcgccccgg

.....686 nt.....

Exon 16 Coordinates: 527d12_Contig309G 47113-47322

ggctgcttcagGGGCCAACCGCCTGACCCTGGAGGACGCCAACATCGTGCAGC
CTCTGGGCCTGACCATCCTTGGCAAGCATCTCTACTGGATCGACCGCCA
GCAGCAGATGATCGAGCGTGTGGAGAAGACCACCGGGGACAAGCGGAC
TCGCATCCAGGGCCGTGTGCGCCACCTCACTGGCATCCATGCAGTGGAG
GAAGTCAGCCTGGAGGAGTTCTgtacgtgggggc

.....3884 nt.....

Exon 17 Coordinates: 527d12_Contig309G 51206-51331

ttgtctttgcagCAGCCCACCCATGTGCCCCTGACAATGGTGGCTGCTCCCACAT
CTGTATTGCCAAGGGTGATGGGACACCACGGTGCTCATGCCCAGTCCAC
CTCGTGCTCCTGCAGAACCTGCTGACCTGTGGAGgtaggtgtgacctaggtgc

....3905 nt.....

Exon 18 Coordinates: 527d12_Contig309G 55236-55472

gttcctctgtccctccccagAGCCGCCCACCTGCTCCCCGGACCAAGTTTGCATGTG
CCACAGGGGAGATCGACTGTATCCCCGGGGCCTGGCGCTGTGACGGCTT
TCCCGAGTGCGATGACCAGAGCGACGAGGAGGGCTGCCCCGTGTGCTCC
GCCGCCCAGTTCCCCTGCGCGCGGGGTCAAGTGTGTGGACCTGCGCCTGC
GCTGCGACGGCGAGGCAGACTGTCAGGACCGCTCAGACGAGGTGGACT
GTGACGgtgaggccctcc

.....3052 nt.....

FIG. 3D

9/29

Exon 19 Coordinates: 527d12_Contig309G 58524-58634

tctccttgagCCATCTGCCTGCCCAACCAGTTCCGGTGTGCGAGCGGCCAGTG
TGTCTCATCAAACAGCAGTGCGACTCCTTCCCCGACTGTATCGACGGCT
CCGACGAGCTCATGTGTGgtgagccagctt

.....1448 nt.....

Exon 20 Coordinates: 527d12_Contig309G 60082-60319

gtttgtcttgagAAATCACCAAGCCGCCCTCAGACGACAGCCCGGCCCCACAGC
AGTGCCATCGGGCCCGTCATTGGCATCATCCTCTCTCTCTTCGTCATGGG
TGGTGTCTATTTTGTGTGCCAGCGCGTGGTGTGCCAGCGCTATGCGGGG
GCCAACGGGGCCCTTCCCGCACGAGTATGTCAGCGGGACCCCGCACGTGC
CCCTCAATTCATAGCCCCGGGCGGTTCCAGCATGGCCCCCTTCACAGgta
aggagcctgagatatggaa

....1095 nt.....

Exon 21 Coordinates: 527d12_Contig309G 61414-61552

cttccttgagGCATCGCATGCGGAAAGTCCATGATGAGCTCCGTGAGCCTGA
TGGGGGGCCGGGGCGGGGTGCCCCCTCTACGACCGGAACACGTCACAG
GGGCCTCGTCCAGCAGCTCGTCCAGCACGAAGGCCACGCTGTACCCGCC
Ggtgaggggagggg

.....6513 nt.....

Exon 22 Coordinates: 527d12_Contig309G 68065-68162

ttggctctcctagATCCTGAACCCGCCGCCCTCCCCGGCCACGGACCCCTCCCT
GTACAACATGGACATGTTCTACTCTTCAAACATTCCGGCCACTGCGAGAC
CGTACAGgtaggacatcccctgag

.....2273 nt.....

FIG. 3E

10/29

Exon 23 Coordinates: 527d12_Contig309G 70435-70901

tcaaacattccggccactgcgagaccgtacagGCCCTACATCATTCGAGGAATGGCGCCCCC
GACGACGCCCTGCAGCACCGACGTGTGTGACAGCGACTACAGCGCCAGC
CGCTGGAAGGCCAGCAAGTACTACCTGGATTTGAACTCGGACTCAGACC
CCTATCCACCCCCACCCACGCCCCACAGCCAGTACCTGTCGGCGGAGGA
CAGCTGCCCCGCCCTCGCCCCGCCACCGAGAGGAGCTACTTCCATCTCTTC
CCGCCCCCTCCGTCCCCCTGCACGGACTCATCCTGACCTCGGCCGGGCCA
CTCTGGCTTCTCTGTGCCCCTGTAAATAGTTTTAAATATGAACAAAGAAAAA
ATATATTTTATGATTTAAAAAATAAATATAATTGGGATTTTAAAAACATGAGA
AATGTGAACTGTGATGGGGTGGGCAGGGCTGGGAGAACTTTGTACAGTGGAG
AAATATTTATAAACTTAATTTGTAAACA

FIG. 3F

11/29

Model for a LDL Receptor-Related protein, Zmax1



FIG. 4

12/29

High Bone Mass Gene

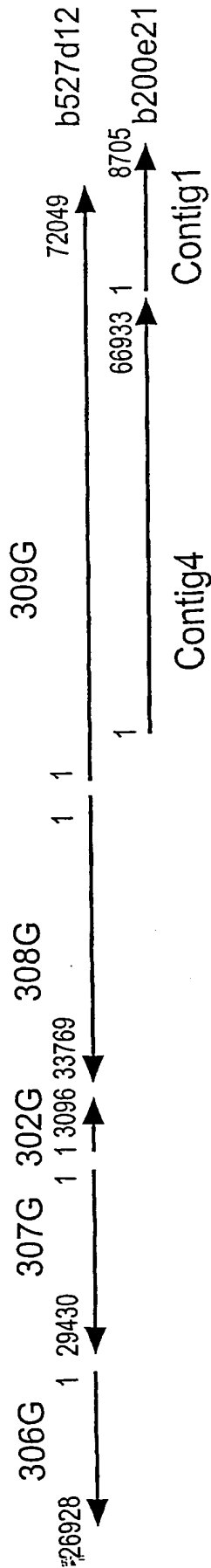


FIG. 5

FIG. 6A

SUBSTITUTE SHEET (RULE 26)

14/29

FIG. 6B

601	GCGGGCAGGGATGGATGGCAGCACCCGGAAGATCATTTGTGACTCGGACATTTACTGGCC	660
178	R A G M D G S T R K I I V D S D I Y W P	197
661	CAATGGACTGACCATCGACCTGGAGGAGCAGAAGCTCTACTGGGCTGACGCCAAGCTCAG	720
198	N G L T I D L E E Q K L Y W A D A K L S	217
721	CTTCATCCACCGTGCCAACTGGACGGCTCGTTCCGGCAGAAGGTGGTGAGGGCAGCCT	780
218	F I H R A N L D G S F R Q K V V E G S L	237
781	GACGCACCCCTTCGCCCTGACGGCTCTCCGGGGACACTCTGTACTGGACAGACTGGCAGAC	840
238	T H P F A L T L S G D T L Y W T D W Q T	257
841	CCGCTCCATCCATGCCCTGCAACAAGCCGCACTGGGGGGAAGAGAGATCCTGAGTGC	900
258	R S I H A C N K R T G G K R K E I L S A	277
901	CCTCTACTACCCATGGACATCCAGGTGCTGAGCCAGGAGCGGCAGCCCTTTCTTCCACAC	960
278	L Y S P M D I Q V L S Q E R Q P F F H T	297
961	TCGCTGTGAGGAGACAATGGCGGTGCTCCACCTGTGCCCTGTGTCCCCAAGCGAGCC	1020
298	R C E E D N G G C S H L C L L S P S E P	317
1021	TTTCTACACATGCGCCTGCCCCACGGGTGTGCAGCTGCAGGACAACGGCAGGACGTGTAA	1080
318	F Y T C A C P T G V Q L Q D N G R T C K	337
1081	GGCAGGAGCGAGGTGCTGTGCTGTGCCCGGCGGACCGACCTACGGAGGATCTCGCT	1140
338	A G A E E V L L L A R R T D L R R I S L	357

15/29

FIG. 6C

1141	GGACACGCCGGA	1200
358	CTTCACCGACATCGTGTG	377
	CTGAGTGACGACATCCGGCACGCCATTGC	
	D T P D F T D I V L Q V D D I R H A I A	
1201	CATCGACTACGACCGCTAGAGGGCTATGTCTACTGGACAGATGACGAGGTGCGGGCCCAT	1260
378	I D Y D P L E G Y V Y W T D D E V R A I	397
1261	CCGACGGGCGTACCTGGACGGGTCTGGGGGGCAGACGCTGGTCAACACCGAGATCAACGA	1320
398	R R A Y L D G S G A Q T L V N T E I N D	417
1321	CCCCGATGCGATCGCGGTGCGACTGGGTGGCCCCGAAACCTCTACTGGACCGACACGGGCAC	1380
418	P D G I A V D W V A R N L Y W T D T G T	437
1381	GGACCGCATCGAGGTGACGCGGCTCAACGGCACCTCCCGCAAGATCCTGGTGTGGAGGA	1440
438	D R I E V T R L N G T S R K I L V S E D	457
1441	CCTGGACGAGCCCCGAGCCCATCGCACTGCACCCCGTGTGATGGGCCCTCATGTACTGGACAGA	1500
458	L D E P R A I A L H P V M G L M Y W T D	477
1501	CTGGGAGAGAACCTAAATCGAGTGTGCCAACTTGGATGGGCAGGACGGCGGTGTGCT	1560
478	W G E N P K I E C A N L D G Q E R R V L	497
1561	GGTCAATGCCCTCCCTCGGGTGGCCCAACGGCCTGGCCCTGGACCTGCAGGAGGGAAGCT	1620
498	V N A S L G W P N G L A L D L Q E G K L	517
1621	CTACTGGGAGACGCCAAGACAGACAAGATCGAGGTGATCAATGTTGATGGGACGAAGAG	1680
518	Y W G D A K T D K I E V I N V D G T K R	537

16/29

FIG. 6D

1681	GCGGACCCCTCCTGGAGGACAAGCTCCCGCACATTTTCGGGTTTCACGCTGCTGGGGACTT	1740
538	R T L L E D K L P H I F G F T L L G D F	557
1741	CATCTACTGGACTGACTGGCAGCGCGCAGCATCGAGCGGTGCACAAGTCAAGGCCAG	1800
558	I Y W T D W Q R R S I E R V H K V K A S	577
1801	CCGGGACGTCAATCATGACCAGCTGCCCGACCTGATGGGGCTCAAAGCTGTGAATGTGC	1860
578	R D V I I D Q L P D L M G L K A V N V A	597
1861	CAAGGTCGTCGGAACCAACCGTGTGCGGACAGGAACGGGGGTGCAGCCACCTGTGCTT	1920
598	K V V G T N P C A D R N G G C S H L C F	617
1921	CTTCACACCCACGCAACCGGTGTGGTGCCCATCGGCCCTGGAGCTGCTGAGTGACAT	1980
618	F T P H A T R C G C P I G L E L L S D M	637
1981	GAAGACCTGCATCGTGCCCTGAGGCCCTTCTTGGTCTTCACCAAGCAGAGCCGCCATCCACAG	2040
638	K T C I V P E A F L V F T S R A A I H R	657
2041	GATCTCCCTCGAGACCAATAACAACGACGTGGCCATCCCGCTCACGGCGTCAAGGAGGC	2100
658	I S L E T N N N D V A I P L T G V K E A	677
2101	CTCAGCCCTGGACTTTGATGTGTCCAAACAACCATCTACTGGACAGACGTCAGCCTGAA	2160
678	S A L D F D V S N N H I Y W T D V S L K	697
2161	GACCATCAGCCGCCCTTCATGAACGGGAGCTCGGTGGAGCACGTTGGAGTTTGGCCT	2220
698	T I S R A F M N G S S V E H V V E F G L	717

17/29

FIG. 6E

2221	TGACTACCCGAGGCGCATGGCCGTTGACTGGGCAAGAACTCTACTGGGCCGACAC	2280
718	D Y P E G M A V D W M G K N L Y W A D T	737
2281	TGGGACCAACAGAAATCGAAGTGGCGCGGCTGACGGGAGTTCCGGCAAGTCCTCGTGTG	2340
738	G T N R I E V A R L D G Q F R Q V L V W	757
2341	GAGGACTTGGACAAACCGAGGTCGCTGGCCCTGGATCCCAAGGGCTACATCTACTG	2400
758	R D L D N P R S L A L D P T K G Y I Y W	777
2401	GACCGAGTGGGGCGCAAGCCGAGGATCGTGGGGCCCTTCATGGACGGGACCAACTGCAT	2460
778	T E W G G K P R I V R A F M D G T N C M	797
2461	GACGCTGGTGGACAAAGTGGCGCGGCGCAACGACCTCACCATTTGACTACGCTGACCAGCG	2520
798	T L V D K V G R A N D L T I D Y A D Q R	817
2521	CCTCTACTGGACCGACCTGGACACCAACATGATCGAGTCGTCCAACATGCTGGGTCAGGA	2580
818	L Y W T D L D T N M I E S S N M L G Q E	837
2581	GCGGTCGTGATTGCCGACGATCTCCCGCACCCGTTCCGGTCTGACGCAGTACAGCGATT	2640
838	R V V I A D D L P H P F G L T Q Y S D Y	857
2641	TATCTACTGGACAGACTGGAATCTGCACAGCATTTGACGGGCGCGCAAGACTAGCGGCCG	2700
858	I Y W T D W N L H S I E R A D K T S G R	877
2701	GAACCGCACCCCTCATCCAGGGCCACCTGGACTTCTGTGATGGACATCCTGGTGTCCACTC	2760
878	N R T L I Q G H L D F V M D I L V F H S	897

18/29

FIG. 6F

2761	CTCCCGCCAGGATGGCCCTCAATGACTGTATGCACAACAACGGGAGTGTTGGGCAGCTGTG	2820
898	S R Q D G L N D C M H N N G Q C G Q L C	917
2821	CCTTGCCATCCCCGGGGCCACCGCTGCGGCTGCGCCTCACACTACACCCTGGACCCAG	2880
918	L A I P G G H R C G C A S H Y T L D P S	937
2881	CAGCCGCAACTGCAGCCCGCCACCCCTTCTTGCTGTTTCAGCCAGAAATCTGCCATCAG	2940
938	S R N C S P P T T F L L F S Q K S A I S	957
2941	TCGGATGATCCCGGACGACGACGACGACCCGGATCTCATCTGCCCCCTGCATGGACTGAG	3000
958	R M I P D D Q H S P D L I L P L H G L R	977
3001	GAACGTCAAAGCCATCGACTATGACCCCACTGGACAAGTTTCATCTACTGGTGGATGGCG	3060
978	N V K A I D Y D P L D K F I Y W V D G R	997
3061	CCAGAACATCAAGCGAGCCCAAGGACGACGGGACCCAGCCCTTTGTTTGACCTCTCTGAG	3120
998	Q N I K R A K D D G T Q P F V L T S L S	1017
3121	CCAAGGCCAAACCCAGACGAGCAGCCCAACGACCTCAGCATCGACATCTACAGCCGGAC	3180
1018	Q G Q N P D R Q P H D L S I D I Y S R T	1037
3181	ACTGTTCTGGACGTGCGAGGCCACCAATACCATCAACGTCCACAGGCTGAGCGGGGAAGC	3240
1038	L F W T C E A T N T I N V H R L S G E A	1057
3241	CATGGGGTGGTGGTGGGACCGCGACAAGCCCGGCGCATCGTCTCAACGCGGA	3300
1058	M G V V L R G D R D K P R A I V V N A E	1077

19/29

FIG. 6G

3301	GCGAGGGTACCTGTACTTCACCAACATGCAGGACCGGGCAGCCAAAGATCGAACGGCAGC	3360
1078	R G Y L Y F T N M Q D R A A K I E R A A	1097
3361	CCTGGACGGCACCGAGCGGAGGTCCTCTTCAACCGGCCCTCATCCGCCCTGTGGCCCT	3420
1098	L D G T E R E V L F T T G L I R P V A L	1117
3421	GGTGGTGACAAACACACTGGGCAAGCTGTTCTGGGTGGACCGGACCTGAAGCGCATTGA	3480
1118	V V D N T L G K L F W V D A D L K R I E	1137
3481	GAGCTGTGACCTGTCAGGGGCCAACCGCCTGACCCCTGGAGGACGCCAACATCGTGCAGCC	3540
1138	S C D L S G A N R L T L E D A N I V Q P	1157
3541	TCCTGGCCCTGACCATCCTTGGCAAGCATCTCTACTGGATCGACCGCCAGCAGCAGATGAT	3600
1158	L G L T I L G K H L Y W I D R Q Q M I	1177
3601	CGAGCGTGTGGAGAAAGACCACCGGGGACAAAGCGGACTCGCATCCAGGGCCGTGTGCCCA	3660
1178	E R V E K T T G D K R T R I Q G R V A H	1197
3661	CCTCACTGGCATCCATGCAGTGGAGGAAGTCAGCCTGGAGGAGTTCTCAGCCACCCATG	3720
1198	L T G I H A V E E V S L E E F S A H P C	1217
3721	TGCCCGTGACAAATGGTGGCTGCTCCACATCTGTATTGCCAAGGGTGATGGACACCACG	3780
1218	A R D N G G C S H I C I A K G D G T P R	1237
3781	GTGCTCATGCCCAGTCCACCTCGTGTCTCTGTCAGAACTGCTGACCTGTGAGAGCCGCC	3840
1238	C S C P V H L V L L Q N L L T C G E P P	1257

20/29

FIG. 6H

3841	CACCTGCTCCCGGACCAAGTTTGCAATGTGCCACAGGGAGATCGACTGTATCCCCGGGC	3900
1258	T C S P D Q Q F A C A T G E I D C I P G A	1277
3901	CTGGCGCTGTGACGGCTTTCGAGTGCATGACAGAGCGACGAGAGGGCTGCCCGGT	3960
1278	W R C D G F P E C D D Q S D E E G C P V	1297
3961	GTGCTCCCGCCCAAGTTCCCTGCGCGGGGTCAAGTGTGGACCTGGCGCTGCGCTG	4020
1298	C S A A Q F P C A R G Q C V D L R L R C	1317
4021	CGACGGCGAGGCAGACTGTCAAGACCGCTCAGACGAGGTGGACTGTGACGCCATCTGCCT	4080
1318	D G E A D C Q D R S D E V D C D A I C L	1337
4081	GCCCAACCAGTTCCGGTGTGCGAGCGGCCAGTGTGTCTCATCAACAGCAGTGCAGTCTC	4140
1338	P N Q F R C A S G Q C V L I K Q Q C D S	1357
4141	CTTCCCCGACTGTATCGACGGCTCCGACGAGCTCATGTGTAAATCACCAGCCGCCCTC	4200
1358	F P D C I D G S D E L M C E I T K P P S	1377
4201	AGACGACAGCCCGCCACAGCAGTGCCATCGGGCCCGTCAATTGGCATCATCTCTCT	4260
1378	D D S P A H S S A I G P V I G I I L S L	1397
4261	CTTCGTCATGGTGTCTATTTGTGTGCCAGCGCGTGGTGTGCCAGCGCTATGCGGG	4320
1398	F V M G G V Y F V C Q R V V C Q R Y A G	1417
4321	GGCCAACGGCCCTTCCCGCAGAGTATGTACGGGGACCCCGCACGTCGCCCTCAATT	4380
1418	A N G P F P H E Y V S G T P H V P L N F	1437

21/29

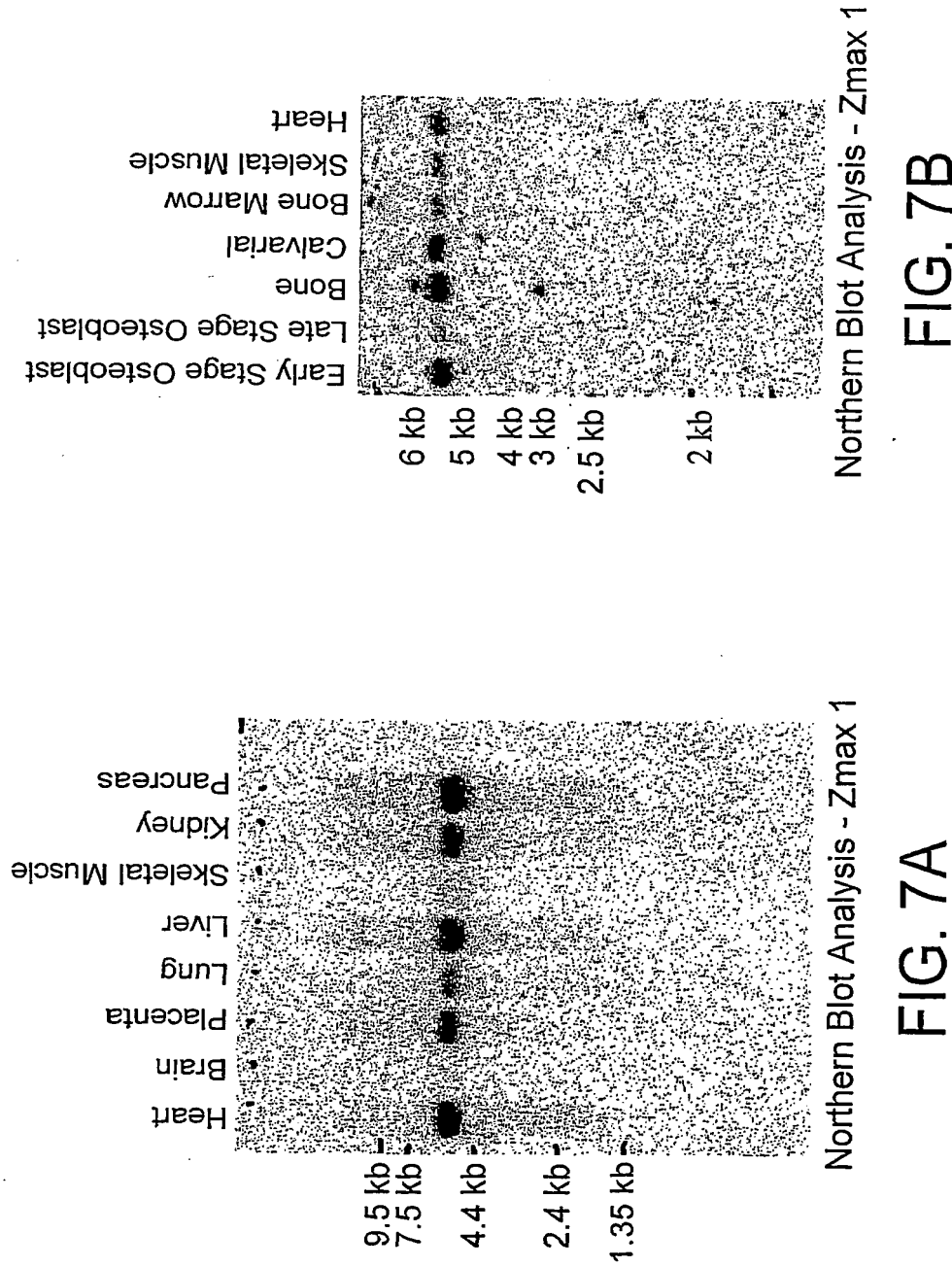
FIG. 6I

4381	CATAGCCCGGGCGGTTCCAGCATGGCCCCCTTACAGGCATCGCGAAAGTCCAT	4440
1438	I A P G G S Q H G P F T G I A C G K S M	1457
4441	GATGAGCTCCGTGAGCCTGATGGGGGGCCGGGGGTGCCCTCTACGACCGGAACCA	4500
1458	M S S V S L M G G R G G V P L Y D R N H	1477
4501	CGTCACAGGGGCCCTCGTCCAGCAGCTCGTCCAGCACGAAGGCCACGCTGTACCCGCCGAT	4560
1478	V T G A S S S S S S T K A T L Y P P I	1497
4561	CCTGAACCCCGCCCTCCCGGCCACGGACCCCTCCCTGTACAACATGGACATGTTCTA	4620
1498	L N P P P S P A T D P S L Y N M D M F Y	1517
4621	CTCTTCAAACATTCGGGCCACTGCGAGACCGTACAGGCCCTACATCATTCGAGGAATGGC	4680
1518	S S N I P A T A R P Y R P Y I I R G M A	1537
4681	GCCCCGACGACGCCCTGCAGCACCGACGTTGTGTACAGCGACTACAGCGCCAGCCGCTG	4740
1538	P P T T P C S T D V C D S D Y S A S R W	1557
4741	GAAGGCCAGCAAGTACTACCTGGATTGAACTCGGACTCAGACCCCTATCCACCCCCACC	4800
1558	K A S K Y Y L D L N S D S D P Y P P P	1577
4801	CACGCCCCACAGCCAGTACCTGTGCGGGGAGGACAGCTGCCCGCCCTCGCCCGCCACCGA	4860
1578	T P H S Q Y L S A E D S C P P S P A T E	1597
4861	GAGGAGCTACTTCCATCTCTTCCCGCCCTCCGTCCCCCTGCACGGACTCATCCTTGACC	4920
1598	R S Y F H L F P P P P S P C T D S S	1615

22/29

FIG. 6J

4921	TCGGCCGGGCCACTCTGGCTTCTCTGTGCCCCCTGTAAATAGTTTAAATATGAACAAGA	4980
4981	AAAAAATATATTTTATGATTTAAAAATAAATAATAATTGGGATTTTAAAAACATGAGAAA	5040
5041	TGTGAAC TGTGATGGGTGGGCAGGGCTGGGAGAAC TTTGTACAGTGGAGAAATATTTAT	5100
5101	AAACTTAA TTTGTAAACA	5120



Zmax 1 random samples

b527d12-h_Contig087C_1.nt

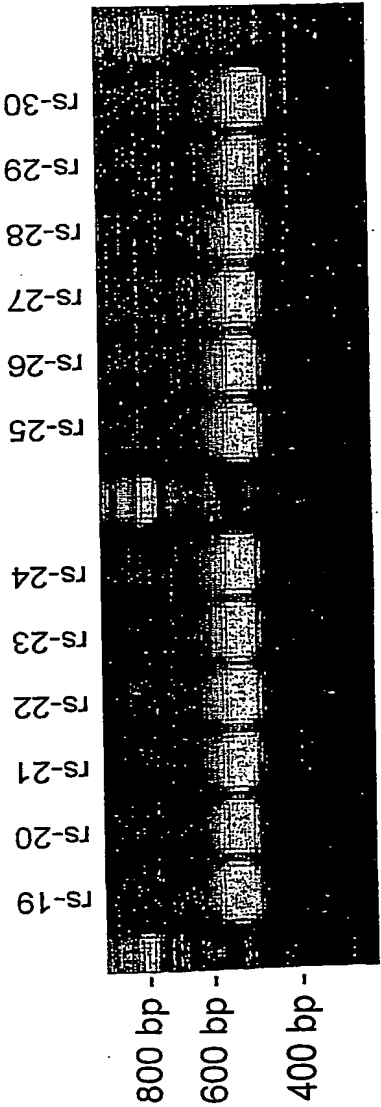


FIG. 8

25/29

ASO Detection of the Zmax1 Exon 3 Mutation

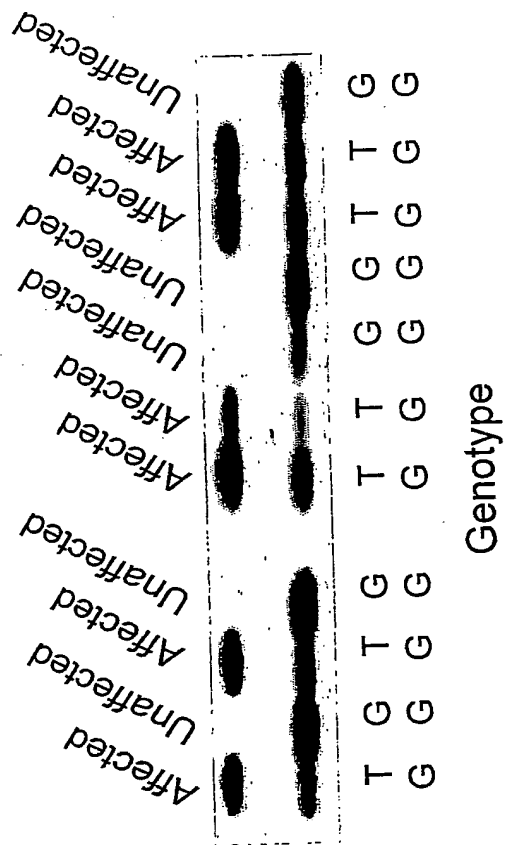


FIG. 9

26/29

Mouse Zmax1 In situ hybridization
100X Magnification

Antisense probe



FIG. 10A

Mouse Zmax1 In situ hybridization
100X Magnification

Sense probe



FIG. 10B

27/29

Mouse Zmax1 In situ hybridization
400X Magnification
Antisense probe

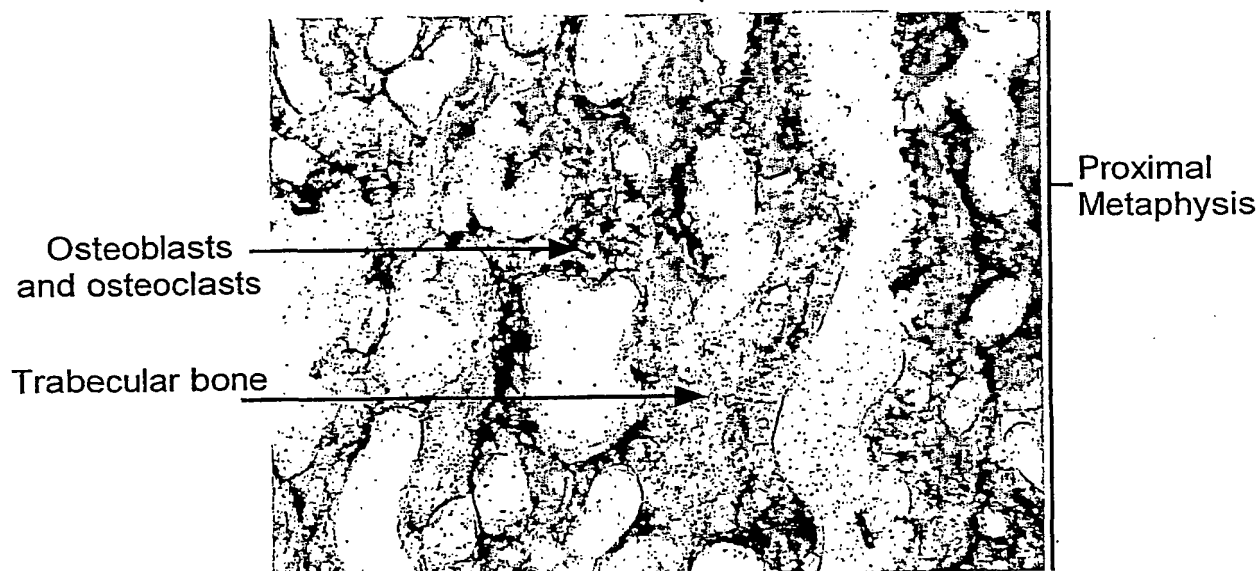


FIG. 11A

Mouse Zmax1 In situ hybridization
400X Magnification
Sense probe

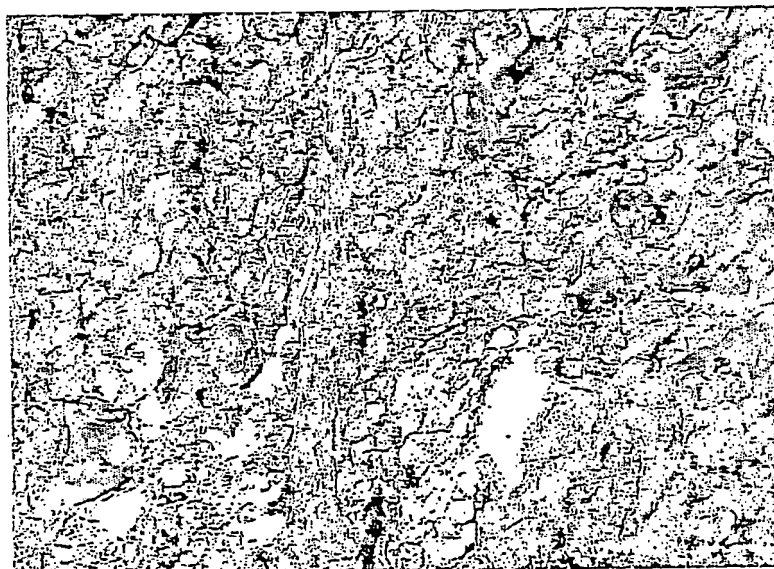
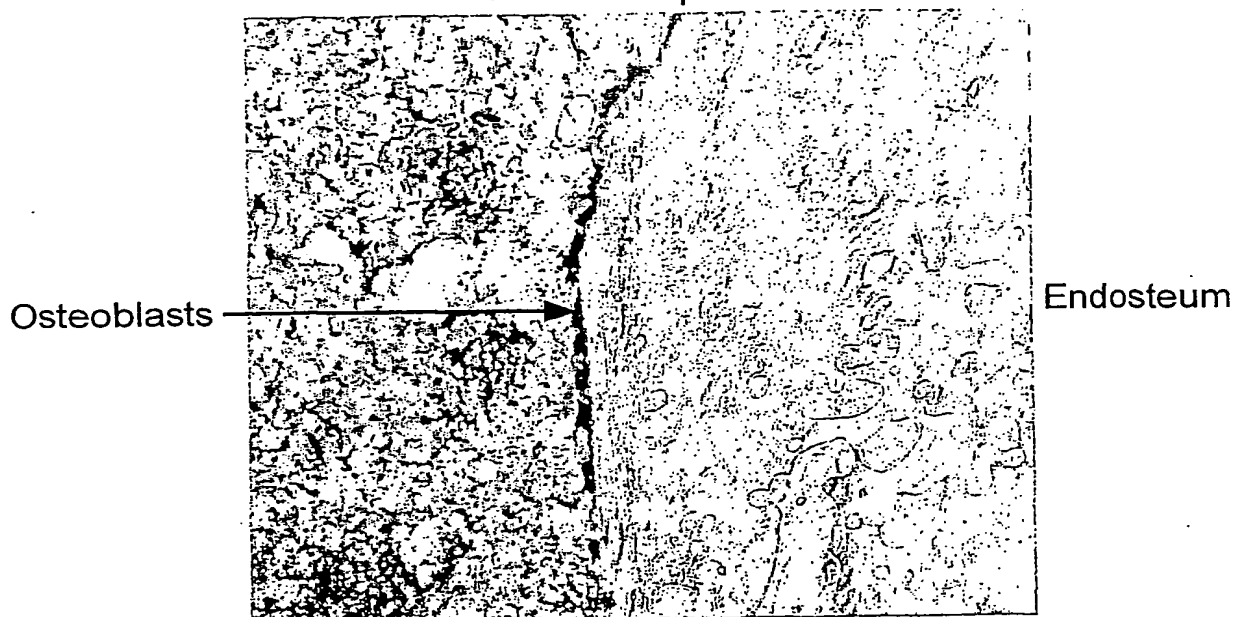


FIG. 11B

28/29

Mouse Zmax1 In situ hybridization
400X Magnification
Antisense probe

**FIG. 12A**

Mouse Zmax1 In situ hybridization
400X Magnification
Sense probe

**FIG. 12B**

SUBSTITUTE SHEET (RULE 26)

29/29

Antisense Inhibition of Zmax1 Expression

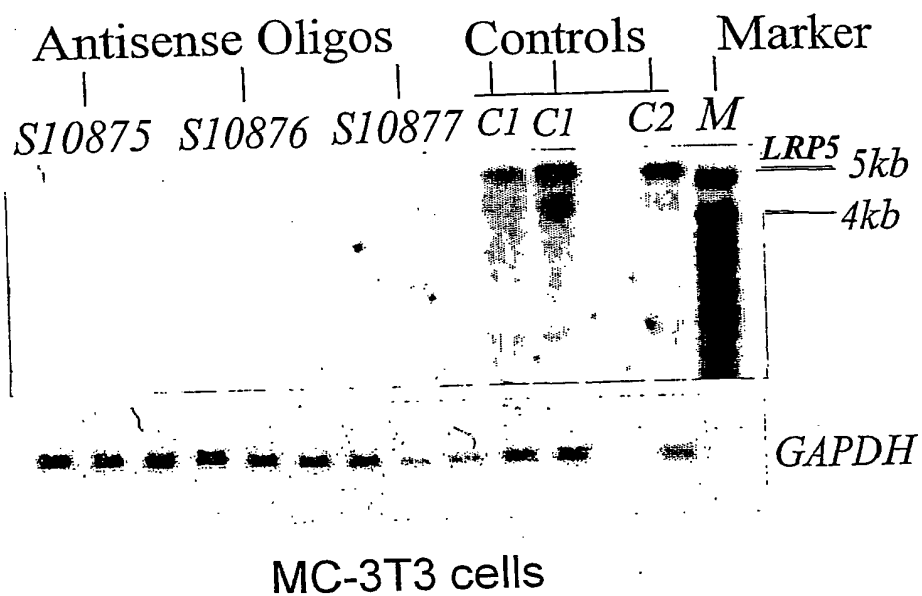


FIG. 13

1 5 10
 ctg ctg ctg ctg ctg ctg gcg ctg tgc ggc tgc ccg gcc ccc gcc 157
 Leu Leu Leu Leu Leu Leu Leu Ala Leu Cys Gly Cys Pro Ala Pro Ala
 15 20 25
 gcg gcc tcg ccg ctc ctg cta ttt gcc aac cgc cgg gac gta cgg ctg 205
 Ala Ala Ser Pro Leu Leu Leu Phe Ala Asn Arg Arg Asp Val Arg Leu
 30 35 40 45
 gtg gac gcc ggc gga gtc aag ctg gag tcc acc atc gtg gtc agc ggc 253
 Val Asp Ala Gly Gly Val Lys Leu Glu Ser Thr Ile Val Val Ser Gly
 50 55 60
 ctg gag gat gcg gcc gca gtg gac ttc cag ttt tcc aag gga gcc gtg 301
 Leu Glu Asp Ala Ala Ala Val Asp Phe Gln Phe Ser Lys Gly Ala Val
 65 70 75
 tac tgg aca gac gtg agc gag gag gcc atc aag cag acc tac ctg aac 349
 Tyr Trp Thr Asp Val Ser Glu Glu Ala Ile Lys Gln Thr Tyr Leu Asn
 80 85 90
 cag acg ggg gcc gcc gtg cag aac gtg gtc atc tcc ggc ctg gtc tct 397
 Gln Thr Gly Ala Ala Val Gln Asn Val Val Ile Ser Gly Leu Val Ser
 95 100 105
 ccc gac ggc ctc gcc tgc gac tgg gtg ggc aag aag ctg tac tgg acg 445
 Pro Asp Gly Leu Ala Cys Asp Trp Val Gly Lys Lys Leu Tyr Trp Thr
 110 115 120 125
 gac tca gag acc aac cgc atc gag gtg gcc aac ctc aat ggc aca tcc 493

SEQUENCE LISTING

<110> John P. Carulli et al.

<120> REGULATING LIPID LEVELS VIA THE ZMAX1 or HBM GENE

<130> 032796-019

<150> Unassigned

<151> 2000-05-26

<150> US 09/543,771

<151> 2000-04-05

<150> US 09/544,398

<151> 2000-04-05

<160> 62

<210> 1

<211> 5120

<212> DNA

<213> Homo sapiens

<400> 1

actaaagcgc cgccgccgcg ccatggagcc cgagtgcgcg cggcgcgggc ccgtccggcc 60

gccggacaac atg gag gca gcg ccg ccc ggg ccg ccg tgg ccg ctg ctg 109

Met Glu Ala Ala Pro Pro Gly Pro Pro Trp Pro Leu Leu

Asp Ser Glu Thr Asn Arg Ile Glu Val Ala Asn Leu Asn Gly Thr Ser

130 135 140

cgg aag gtg ctc ttc tgg cag gac ctt gac cag ccg agg gcc atc gcc 541

Arg Lys Val Leu Phe Trp Gln Asp Leu Asp Gln Pro Arg Ala Ile Ala

145 150 155

ttg gac ccc gct cac ggg tac atg tac tgg aca gac tgg ggt gag acg 589

Leu Asp Pro Ala His Gly Tyr Met Tyr Trp Thr Asp Trp Gly Glu Thr

160 165 170

ccc cgg att gag cgg gca ggg atg gat ggc agc acc cgg aag atc att 637

Pro Arg Ile Glu Arg Ala Gly Met Asp Gly Ser Thr Arg Lys Ile Ile

175 180 185

gtg gac tcg gac att tac tgg ccc aat gga ctg acc atc gac ctg gag 685

Val Asp Ser Asp Ile Tyr Trp Pro Asn Gly Leu Thr Ile Asp Leu Glu

190 195 200 205

gag cag aag ctc tac tgg gct gac gcc aag ctc agc ttc atc cac cgt 733

Glu Gln Lys Leu Tyr Trp Ala Asp Ala Lys Leu Ser Phe Ile His Arg

210 215 220

gcc aac ctg gac ggc tcg ttc cgg cag aag gtg gtg gag ggc agc ctg 781

Ala Asn Leu Asp Gly Ser Phe Arg Gln Lys Val Val Glu Gly Ser Leu

225 230 235

acg cac ccc ttc gcc ctg acg ctc tcc ggg gac act ctg tac tgg aca 829

Thr His Pro Phe Ala Leu Thr Leu Ser Gly Asp Thr Leu Tyr Trp Thr

240 245 250

gac tgg cag acc cgc tcc atc cat gcc tgc aac aag cgc act ggg ggg 877
Asp Trp Gln Thr Arg Ser Ile His Ala Cys Asn Lys Arg Thr Gly Gly
255 260 265
aag agg aag gag atc ctg agt gcc ctc tac tca ccc atg gac atc cag 925
Lys Arg Lys Glu Ile Leu Ser Ala Leu Tyr Ser Pro Met Asp Ile Gln
270 275 280 285
gtg ctg agc cag gag cgg cag cct ttc ttc cac act cgc tgt gag gag 973
Val Leu Ser Gln Glu Arg Gln Pro Phe Phe His Thr Arg Cys Glu Glu
290 295 300
gac aat ggc ggc tgc tcc cac ctg tgc ctg ctg tcc cca agc gag cct 1021
Asp Asn Gly Gly Cys Ser His Leu Cys Leu Leu Ser Pro Ser Glu Pro
305 310 315
ttc tac aca tgc gcc tgc ccc acg ggt gtg cag ctg cag gac aac ggc 1069
Phe Tyr Thr Cys Ala Cys Pro Thr Gly Val Gln Leu Gln Asp Asn Gly
320 325 330
agg acg tgt aag gca gga gcc gag gag gtg ctg ctg ctg gcc cgg cgg 1117
Arg Thr Cys Lys Ala Gly Ala Glu Glu Val Leu Leu Leu Ala Arg Arg
335 340 345
acg gac cta cgg agg atc tcg ctg gac acg ccg gac ttc acc gac atc 1165
Thr Asp Leu Arg Arg Ile Ser Leu Asp Thr Pro Asp Phe Thr Asp Ile
350 355 360 365
gtg ctg cag gtg gac gac atc cgg cac gcc att gcc atc gac tac gac 1213
Val Leu Gln Val Asp Asp Ile Arg His Ala Ile Ala Ile Asp Tyr Asp

370 375 380
ccg cta gag ggc tat gtc tac tgg aca gat gac gag gtg cgg gcc atc 1261
Pro Leu Glu Gly Tyr Val Tyr Trp Thr Asp Asp Glu Val Arg Ala Ile
385 390 395
cgc agg gcg tac ctg gac ggg tct ggg gcg cag acg ctg gtc aac acc 1309
Arg Arg Ala Tyr Leu Asp Gly Ser Gly Ala Gln Thr Leu Val Asn Thr
400 405 410
gag atc aac gac ccc gat ggc atc gcg gtc gac tgg gtg gcc cga aac 1357
Glu Ile Asn Asp Pro Asp Gly Ile Ala Val Asp Trp Val Ala Arg Asn
415 420 425
ctc tac tgg acc gac acg ggc acg gac cgc atc gag gtg acg cgc ctc 1405
Leu Tyr Trp Thr Asp Thr Gly Thr Asp Arg Ile Glu Val Thr Arg Leu
430 435 440 445
aac ggc acc tcc cgc aag atc ctg gtg tcg gag gac ctg gac gag ccc 1453
Asn Gly Thr Ser Arg Lys Ile Leu Val Ser Glu Asp Leu Asp Glu Pro
450 455 460
cga gcc atc gca ctg cac ccc gtg atg ggc ctc atg tac tgg aca gac 1501
Arg Ala Ile Ala Leu His Pro Val Met Gly Leu Met Tyr Trp Thr Asp
465 470 475
tgg gga gag aac cct aaa atc gag tgt gcc aac ttg gat ggg cag gag 1549
Trp Gly Glu Asn Pro Lys Ile Glu Cys Ala Asn Leu Asp Gly Gln Glu
480 485 490
cgg cgt gtg ctg gtc aat gcc tcc ctc ggg tgg ccc aac ggc ctg gcc 1597

Arg Arg Val Leu Val Asn Ala Ser Leu Gly Trp Pro Asn Gly Leu Ala
495 500 505
ctg gac ctg cag gag ggg aag ctc tac tgg gga gac gcc aag aca gac 1645
Leu Asp Leu Gln Glu Gly Lys Leu Tyr Trp Gly Asp Ala Lys Thr Asp
510 515 520 525
aag atc gag gtg atc aat gtt gat ggg acg aag agg cgg acc ctc ctg 1693
Lys Ile Glu Val Ile Asn Val Asp Gly Thr Lys Arg Arg Thr Leu Leu
530 535 540
gag gac aag ctc ccg cac att ttc ggg ttc acg ctg ctg ggg gac ttc 1741
Glu Asp Lys Leu Pro His Ile Phe Gly Phe Thr Leu Leu Gly Asp Phe
545 550 555
atc tac tgg act gac tgg cag cgc cgc agc atc gag cgg gtg cac aag 1789
Ile Tyr Trp Thr Asp Trp Gln Arg Arg Ser Ile Glu Arg Val His Lys
560 565 570
gtc aag gcc agc cgg gac gtc atc att gac cag ctg ccc gac ctg atg 1837
Val Lys Ala Ser Arg Asp Val Ile Ile Asp Gln Leu Pro Asp Leu Met
575 580 585
ggg ctc aaa gct gtg aat gtg gcc aag gtc gtc gga acc aac ccg tgt 1885
Gly Leu Lys Ala Val Asn Val Ala Lys Val Val Gly Thr Asn Pro Cys
590 595 600 605
gcg gac agg aac ggg ggg tgc agc cac ctg tgc ttc ttc aca ccc cac 1933
Ala Asp Arg Asn Gly Gly Cys Ser His Leu Cys Phe Phe Thr Pro His
610 615 620

gca acc cgg tgt ggc tgc ccc atc ggc ctg gag ctg ctg agt gac atg 1981

Ala Thr Arg Cys Gly Cys Pro Ile Gly Leu Glu Leu Leu Ser Asp Met

625

630

635

aag acc tgc atc gtg cct gag gcc ttc ttg gtc ttc acc agc aga gcc 2029

Lys Thr Cys Ile Val Pro Glu Ala Phe Leu Val Phe Thr Ser Arg Ala

640

645

650

gcc atc cac agg atc tcc ctc gag acc aat aac aac gac gtg gcc atc 2077

Ala Ile His Arg Ile Ser Leu Glu Thr Asn Asn Asn Asp Val Ala Ile

655

660

665

ccg ctc acg ggc gtc aag gag gcc tca gcc ctg gac ttt gat gtg tcc 2125

Pro Leu Thr Gly Val Lys Glu Ala Ser Ala Leu Asp Phe Asp Val Ser

670

675

680

685

aac aac cac atc tac tgg aca gac gtc agc ctg aag acc atc agc cgc 2173

Asn Asn His Ile Tyr Trp Thr Asp Val Ser Leu Lys Thr Ile Ser Arg

690

695

700

gcc ttc atg aac ggg agc tcg gtg gag cac gtg gtg gag ttt ggc ctt 2221

Ala Phe Met Asn Gly Ser Ser Val Glu His Val Val Glu Phe Gly Leu

705

710

715

gac tac ccc gag ggc atg gcc gtt gac tgg atg ggc aag aac ctc tac 2269

Asp Tyr Pro Glu Gly Met Ala Val Asp Trp Met Gly Lys Asn Leu Tyr

720

725

730

tgg gcc gac act ggg acc aac aga atc gaa gtg gcg cgg ctg gac ggg 2317

Trp Ala Asp Thr Gly Thr Asn Arg Ile Glu Val Ala Arg Leu Asp Gly

735 740 745
 cag ttc cgg caa gtc ctc gtg tgg agg gac ttg gac aac ccg agg tcg 2365
 Gln Phe Arg Gln Val Leu Val Trp Arg Asp Leu Asp Asn Pro Arg Ser
 750 755 760 765
 ctg gcc ctg gat ccc acc aag ggc tac atc tac tgg acc gag tgg ggc 2413
 Leu Ala Leu Asp Pro Thr Lys Gly Tyr Ile Tyr Trp Thr Glu Trp Gly
 770 775 780
 ggc aag ccg agg atc gtg cgg gcc ttc atg gac ggg acc aac tgc atg 2461
 Gly Lys Pro Arg Ile Val Arg Ala Phe Met Asp Gly Thr Asn Cys Met
 785 790 795
 acg ctg gtg gac aag gtg ggc cgg gcc aac gac ctc acc att gac tac 2509
 Thr Leu Val Asp Lys Val Gly Arg Ala Asn Asp Leu Thr Ile Asp Tyr
 800 805 810
 gct gac cag cgc ctc tac tgg acc gac ctg gac acc aac atg atc gag 2557
 Ala Asp Gln Arg Leu Tyr Trp Thr Asp Leu Asp Thr Asn Met Ile Glu
 815 820 825
 tcg tcc aac atg ctg ggt cag gag cgg gtc gtg att gcc gac gat ctc 2605
 Ser Ser Asn Met Leu Gly Gln Glu Arg Val Val Ile Ala Asp Asp Leu
 830 835 840 845
 ccg cac ccg ttc ggt ctg acg cag tac agc gat tat atc tac tgg aca 2653
 Pro His Pro Phe Gly Leu Thr Gln Tyr Ser Asp Tyr Ile Tyr Trp Thr
 850 855 860
 gac tgg aat ctg cac agc att gag cgg gcc gac aag act agc ggc cgg 2701

Asp Trp Asn Leu His Ser Ile Glu Arg Ala Asp Lys Thr Ser Gly Arg

865 870 875

aac cgc acc ctc atc cag ggc cac ctg gac ttc gtg atg gac atc ctg 2749

Asn Arg Thr Leu Ile Gln Gly His Leu Asp Phe Val Met Asp Ile Leu

880 885 890

gtg ttc cac tcc tcc cgc cag gat ggc ctc aat gac tgt atg cac aac 2797

Val Phe His Ser Ser Arg Gln Asp Gly Leu Asn Asp Cys Met His Asn

895 900 905

aac ggg cag tgt ggg cag ctg tgc ctt gcc atc ccc ggc ggc cac cgc 2845

Asn Gly Gln Cys Gly Gln Leu Cys Leu Ala Ile Pro Gly Gly His Arg

910 915 920 925

tgc ggc tgc gcc tca cac tac acc ctg gac ccc agc agc cgc aac tgc 2893

Cys Gly Cys Ala Ser His Tyr Thr Leu Asp Pro Ser Ser Arg Asn Cys

930 935 940

agc ccg ccc acc acc ttc ttg ctg ttc agc cag aaa tct gcc atc agt 2941

Ser Pro Pro Thr Thr Phe Leu Leu Phe Ser Gln Lys Ser Ala Ile Ser

945 950 955

cgg atg atc ccg gac gac cag cac agc ccg gat ctc atc ctg ccc ctg 2989

Arg Met Ile Pro Asp Asp Gln His Ser Pro Asp Leu Ile Leu Pro Leu

960 965 970

cat gga ctg agg aac gtc aaa gcc atc gac tat gac cca ctg gac aag 3037

His Gly Leu Arg Asn Val Lys Ala Ile Asp Tyr Asp Pro Leu Asp Lys

975 980 985

ttc atc tac tgg gtg gat ggg cgc cag aac atc aag cga gcc aag gac 3085
Phe Ile Tyr Trp Val Asp Gly Arg Gln Asn Ile Lys Arg Ala Lys Asp
990 995 1000 1005
gac ggg acc cag ccc ttt gtt ttg acc tct ctg agc caa ggc caa aac 3133
Asp Gly Thr Gln Pro Phe Val Leu Thr Ser Leu Ser Gln Gly Gln Asn
1010 1015 1020
cca gac agg cag ccc cac gac ctc agc atc gac atc tac agc cgg aca 3181
Pro Asp Arg Gln Pro His Asp Leu Ser Ile Asp Ile Tyr Ser Arg Thr
1025 1030 1035
ctg ttc tgg acg tgc gag gcc acc aat acc atc aac gtc cac agg ctg 3229
Leu Phe Trp Thr Cys Glu Ala Thr Asn Thr Ile Asn Val His Arg Leu
1040 1045 1050
agc ggg gaa gcc atg ggg gtg gtg ctg cgt ggg gac cgc gac aag ccc 3277
Ser Gly Glu Ala Met Gly Val Val Leu Arg Gly Asp Arg Asp Lys Pro
1055 1060 1065
agg gcc atc gtc gtc aac gcg gag cga ggg tac ctg tac ttc acc aac 3325
Arg Ala Ile Val Val Asn Ala Glu Arg Gly Tyr Leu Tyr Phe Thr Asn
1070 1075 1080 1085
atg cag gac cgg gca gcc aag atc gaa cgc gca gcc ctg gac ggc acc 3373
Met Gln Asp Arg Ala Ala Lys Ile Glu Arg Ala Ala Leu Asp Gly Thr
1090 1095 1100
gag cgc gag gtc ctc ttc acc acc ggc ctc atc cgc cct gtg gcc ctg 3421
Glu Arg Glu Val Leu Phe Thr Thr Gly Leu Ile Arg Pro Val Ala Leu

1105 1110 1115
gtg gtg gac aac aca ctg ggc aag ctg ttc tgg gtg gac gcg gac ctg 3469
Val Val Asp Asn Thr Leu Gly Lys Leu Phe Trp Val Asp Ala Asp Leu
1120 1125 1130
aag cgc att gag agc tgt gac ctg tca ggg gcc aac cgc ctg acc ctg 3517
Lys Arg Ile Glu Ser Cys Asp Leu Ser Gly Ala Asn Arg Leu Thr Leu
1135 1140 1145
gag gac gcc aac atc gtg cag cct ctg ggc ctg acc atc ctt ggc aag 3565
Glu Asp Ala Asn Ile Val Gln Pro Leu Gly Leu Thr Ile Leu Gly Lys
1150 1155 1160 1165
cat ctc tac tgg atc gac cgc cag cag cag atg atc gag cgt gtg gag 3613
His Leu Tyr Trp Ile Asp Arg Gln Gln Gln Met Ile Glu Arg Val Glu
1170 1175 1180
aag acc acc ggg gac aag cgg act cgc atc cag ggc cgt gtc gcc cac 3661
Lys Thr Thr Gly Asp Lys Arg Thr Arg Ile Gln Gly Arg Val Ala His
1185 1190 1195
ctc act ggc atc cat gca gtg gag gaa gtc agc ctg gag gag ttc tca 3709
Leu Thr Gly Ile His Ala Val Glu Glu Val Ser Leu Glu Glu Phe Ser
1200 1205 1210
gcc cac cca tgt gcc cgt gac aat ggt ggc tgc tcc cac atc tgt att 3757
Ala His Pro Cys Ala Arg Asp Asn Gly Gly Cys Ser His Ile Cys Ile
1215 1220 1225
gcc aag ggt gat ggg aca cca cgg tgc tca tgc cca gtc cac ctc gtg 3805

Ala Lys Gly Asp Gly Thr Pro Arg Cys Ser Cys Pro Val His Leu Val

1230 1235 1240 1245

ctc ctg cag aac ctg ctg acc tgt gga gag ccg ccc acc tgc tcc ccg 3853

Leu Leu Gln Asn Leu Leu Thr Cys Gly Glu Pro Pro Thr Cys Ser Pro

1250 1255 1260

gac cag ttt gca tgt gcc aca ggg gag atc gac tgt atc ccc ggg gcc 3901

Asp Gln Phe Ala Cys Ala Thr Gly Glu Ile Asp Cys Ile Pro Gly Ala

1265 1270 1275

tgg cgc tgt gac ggc ttt ccc gag tgc gat gac cag agc gac gag gag 3949

Trp Arg Cys Asp Gly Phe Pro Glu Cys Asp Asp Gln Ser Asp Glu Glu

1280 1285 1290

ggc tgc ccc gtg tgc tcc gcc gcc cag ttc ccc tgc gcg cgg ggt cag 3997

Gly Cys Pro Val Cys Ser Ala Ala Gln Phe Pro Cys Ala Arg Gly Gln

1295 1300 1305

tgt gtg gac ctg cgc ctg cgc tgc gac ggc gag gca gac tgt cag gac 4045

Cys Val Asp Leu Arg Leu Arg Cys Asp Gly Glu Ala Asp Cys Gln Asp

1310 1315 1320 1325

cgc tca gac gag gtg gac tgt gac gcc atc tgc ctg ccc aac cag ttc 4093

Arg Ser Asp Glu Val Asp Cys Asp Ala Ile Cys Leu Pro Asn Gln Phe

1330 1335 1340

cgg tgt gcg agc ggc cag tgt gtc ctc atc aaa cag cag tgc gac tcc 4141

Arg Cys Ala Ser Gly Gln Cys Val Leu Ile Lys Gln Gln Cys Asp Ser

1345 1350 1355

ttc ccc gac tgt atc gac ggc tcc gac gag ctc atg tgt gaa atc acc 4189
Phe Pro Asp Cys Ile Asp Gly Ser Asp Glu Leu Met Cys Glu Ile Thr
1360 1365 1370

aag ccg ccc tca gac gac agc ccg gcc cac agc agt gcc atc ggg ccc 4237
Lys Pro Pro Ser Asp Asp Ser Pro Ala His Ser Ser Ala Ile Gly Pro
1375 1380 1385

gtc att ggc atc atc ctc tct ctc ttc gtc atg ggt ggt gtc tat ttt 4285
Val Ile Gly Ile Ile Leu Ser Leu Phe Val Met Gly Gly Val Tyr Phe
1390 1395 1400 1405

gtg tgc cag cgc gtg gtg tgc cag cgc tat gcg ggg gcc aac ggg ccc 4333
Val Cys Gln Arg Val Val Cys Gln Arg Tyr Ala Gly Ala Asn Gly Pro
1410 1415 1420

ttc ccg cac gag tat gtc agc ggg acc ccg cac gtg ccc ctc aat ttc 4381
Phe Pro His Glu Tyr Val Ser Gly Thr Pro His Val Pro Leu Asn Phe
1425 1430 1435

ata gcc ccg ggc ggt tcc cag cat ggc ccc ttc aca ggc atc gca tgc 4429
Ile Ala Pro Gly Gly Ser Gln His Gly Pro Phe Thr Gly Ile Ala Cys
1440 1445 1450

gga aag tcc atg atg agc tcc gtg agc ctg atg ggg ggc cgg ggc ggg 4477
Gly Lys Ser Met Met Ser Ser Val Ser Leu Met Gly Gly Arg Gly Gly
1455 1460 1465

gtg ccc ctc tac gac cgg aac cac gtc aca ggg gcc tcg tcc agc agc 4525
Val Pro Leu Tyr Asp Arg Asn His Val Thr Gly Ala Ser Ser Ser Ser

1470 1475 1480 1485
tcg tcc agc acg aag gcc acg ctg tac ccg ccg atc ctg aac ccg ccg 4573
Ser Ser Ser Thr Lys Ala Thr Leu Tyr Pro Pro Ile Leu Asn Pro Pro
1490 1495 1500
ccc tcc ccg gcc acg gac ccc tcc ctg tac aac atg gac atg ttc tac 4621
Pro Ser Pro Ala Thr Asp Pro Ser Leu Tyr Asn Met Asp Met Phe Tyr
1505 1510 1515
tct tca aac att ccg gcc act gcg aga ccg tac agg ccc tac atc att 4669
Ser Ser Asn Ile Pro Ala Thr Ala Arg Pro Tyr Arg Pro Tyr Ile Ile
1520 1525 1530
cga gga atg gcg ccc ccg acg acg ccc tgc agc acc gac gtg tgt gac 4717
Arg Gly Met Ala Pro Pro Thr Thr Pro Cys Ser Thr Asp Val Cys Asp
1535 1540 1545
agc gac tac agc gcc agc cgc tgg aag gcc agc aag tac tac ctg gat 4765
Ser Asp Tyr Ser Ala Ser Arg Trp Lys Ala Ser Lys Tyr Tyr Leu Asp
1550 1555 1560 1565
ttg aac tcg gac tca gac ccc tat cca ccc cca ccc acg ccc cac agc 4813
Leu Asn Ser Asp Ser Asp Pro Tyr Pro Pro Pro Pro Thr Pro His Ser
1570 1575 1580
cag tac ctg tcg gcg gag gac agc tgc ccg ccc tcg ccc gcc acc gag 4861
Gln Tyr Leu Ser Ala Glu Asp Ser Cys Pro Pro Ser Pro Ala Thr Glu
1585 1590 1595
agg agc tac ttc cat ctc ttc ccg ccc cct ccg tcc ccc tgc acg gac 4909

Arg Ser Tyr Phe His Leu Phe Pro Pro Pro Pro Ser Pro Cys Thr Asp

1600 1605 1610

tca tcc tgacctcggc cgggccactc tggcttctct gtgccctgt aaatagttt 4965

Ser Ser

1615

aaatatgaac aaagaaaaaa atatatttta tgatttaaaa aataaatata attgggattt 5025

taaaaacatg agaaatgtga actgtgatgg ggtgggcagg gctgggagaa cttgtacag 5085

tggagaaata ttataaaact taattttgta aaaca 5120

<210> 2

<211> 5120

<212> DNA

<213> Homo sapiens

<400> 2

actaaagcgc cgccgccgcg ccatggagcc cgagtgcgcg cggcgcgggc ccgtccggcc 60

gccggacaac atg gag gca gcg ccg ccc ggg ccg ccg tgg ccg ctg ctg 109

Met Glu Ala Ala Pro Pro Gly Pro Pro Trp Pro Leu Leu

1 5 10

ctg ctg ctg ctg ctg ctg gcg ctg tgc ggc tgc ccg gcc ccc gcc 157

Leu Leu Leu Leu Leu Leu Ala Leu Cys Gly Cys Pro Ala Pro Ala

15 20 25

gcg gcc tcg ccg ctc ctg cta ttt gcc aac cgc cgg gac gta cgg ctg 205

Ala Ala Ser Pro Leu Leu Leu Phe Ala Asn Arg Arg Asp Val Arg Leu

30 35 40 45

gtg gac gcc ggc gga gtc aag ctg gag tcc acc atc gtg gtc agc ggc 253

Val Asp Ala Gly Gly Val Lys Leu Glu Ser Thr Ile Val Val Ser Gly

50 55 60

ctg gag gat gcg gcc gca gtg gac ttc cag ttt tcc aag gga gcc gtg 301

Leu Glu Asp Ala Ala Ala Val Asp Phe Gln Phe Ser Lys Gly Ala Val

65 70 75

tac tgg aca gac gtg agc gag gag gcc atc aag cag acc tac ctg aac 349

Tyr Trp Thr Asp Val Ser Glu Glu Ala Ile Lys Gln Thr Tyr Leu Asn

80 85 90

cag acg ggg gcc gcc gtg cag aac gtg gtc atc tcc ggc ctg gtc tct 397

Gln Thr Gly Ala Ala Val Gln Asn Val Val Ile Ser Gly Leu Val Ser

95 100 105

ccc gac ggc ctc gcc tgc gac tgg gtg ggc aag aag ctg tac tgg acg 445

Pro Asp Gly Leu Ala Cys Asp Trp Val Gly Lys Lys Leu Tyr Trp Thr

110 115 120 125

gac tca gag acc aac cgc atc gag gtg gcc aac ctc aat ggc aca tcc 493

Asp Ser Glu Thr Asn Arg Ile Glu Val Ala Asn Leu Asn Gly Thr Ser

130 135 140

cgg aag gtg ctc ttc tgg cag gac ctt gac cag ccg agg gcc atc gcc 541

Arg Lys Val Leu Phe Trp Gln Asp Leu Asp Gln Pro Arg Ala Ile Ala

145 150 155

ttg gac ccc gct cac ggg tac atg tac tgg aca gac tgg gtt gag acg 589

Leu Asp Pro Ala His Gly Tyr Met Tyr Trp Thr Asp Trp Val Glu Thr

160

165

170

ccc cgg att gag cgg gca ggg atg gat ggc agc acc cgg aag atc att 637

Pro Arg Ile Glu Arg Ala Gly Met Asp Gly Ser Thr Arg Lys Ile Ile

175

180

185

gtg gac tcg gac att tac tgg ccc aat gga ctg acc atc gac ctg gag 685

Val Asp Ser Asp Ile Tyr Trp Pro Asn Gly Leu Thr Ile Asp Leu Glu

190

195

200

205

gag cag aag ctc tac tgg gct gac gcc aag ctc agc ttc atc cac cgt 733

Glu Gln Lys Leu Tyr Trp Ala Asp Ala Lys Leu Ser Phe Ile His Arg

210

215

220

gcc aac ctg gac ggc tcg ttc cgg cag aag gtg gtg gag ggc agc ctg 781

Ala Asn Leu Asp Gly Ser Phe Arg Gln Lys Val Val Glu Gly Ser Leu

225

230

235

acg cac ccc ttc gcc ctg acg ctc tcc ggg gac act ctg tac tgg aca 829

Thr His Pro Phe Ala Leu Thr Leu Ser Gly Asp Thr Leu Tyr Trp Thr

240

245

250

gac tgg cag acc cgc tcc atc cat gcc tgc aac aag cgc act ggg ggg 877

Asp Trp Gln Thr Arg Ser Ile His Ala Cys Asn Lys Arg Thr Gly Gly

255

260

265

aag agg aag gag atc ctg agt gcc ctc tac tca ccc atg gac atc cag 925

Lys Arg Lys Glu Ile Leu Ser Ala Leu Tyr Ser Pro Met Asp Ile Gln

270 275 280 285
gtg ctg agc cag gag cgg cag cct ttc ttc cac act cgc tgt gag gag 973
Val Leu Ser Gln Glu Arg Gln Pro Phe Phe His Thr Arg Cys Glu Glu
 290 295 300
gac aat ggc ggc tgc tcc cac ctg tgc ctg ctg tcc cca agc gag cct 1021
Asp Asn Gly Gly Cys Ser His Leu Cys Leu Leu Ser Pro Ser Glu Pro
 305 310 315
ttc tac aca tgc gcc tgc ccc acg ggt gtg cag ctg cag gac aac ggc 1069
Phe Tyr Thr Cys Ala Cys Pro Thr Gly Val Gln Leu Gln Asp Asn Gly
 320 325 330
agg acg tgt aag gca gga gcc gag gag gtg ctg ctg ctg gcc cgg cgg 1117
Arg Thr Cys Lys Ala Gly Ala Glu Glu Val Leu Leu Leu Ala Arg Arg
 335 340 345
acg gac cta cgg agg atc tcg ctg gac acg ccg gac ttc acc gac atc 1165
Thr Asp Leu Arg Arg Ile Ser Leu Asp Thr Pro Asp Phe Thr Asp Ile
350 355 360 365
gtg ctg cag gtg gac gac atc cgg cac gcc att gcc atc gac tac gac 1213
Val Leu Gln Val Asp Asp Ile Arg His Ala Ile Ala Ile Asp Tyr Asp
 370 375 380
ccg cta gag ggc tat gtc tac tgg aca gat gac gag gtg cgg gcc atc 1261
Pro Leu Glu Gly Tyr Val Tyr Trp Thr Asp Asp Glu Val Arg Ala Ile
 385 390 395
cgc agg gcg tac ctg gac ggg tct ggg gcg cag acg ctg gtc aac acc 1309

Arg Arg Ala Tyr Leu Asp Gly Ser Gly Ala Gln Thr Leu Val Asn Thr

400 405 410

gag atc aac gac ccc gat ggc atc gcg gtc gac tgg gtg gcc cga aac 1357

Glu Ile Asn Asp Pro Asp Gly Ile Ala Val Asp Trp Val Ala Arg Asn

415 420 425

ctc tac tgg acc gac acg ggc acg gac cgc atc gag gtg acg cgc ctc 1405

Leu Tyr Trp Thr Asp Thr Gly Thr Asp Arg Ile Glu Val Thr Arg Leu

430 435 440 445

aac ggc acc tcc cgc aag atc ctg gtg tcg gag gac ctg gac gag ccc 1453

Asn Gly Thr Ser Arg Lys Ile Leu Val Ser Glu Asp Leu Asp Glu Pro

450 455 460

cga gcc atc gca ctg cac ccc gtg atg ggc ctc atg tac tgg aca gac 1501

Arg Ala Ile Ala Leu His Pro Val Met Gly Leu Met Tyr Trp Thr Asp

465 470 475

tgg gga gag aac cct aaa atc gag tgt gcc aac ttg gat ggg cag gag 1549

Trp Gly Glu Asn Pro Lys Ile Glu Cys Ala Asn Leu Asp Gly Gln Glu

480 485 490

cgg cgt gtg ctg gtc aat gcc tcc ctc ggg tgg ccc aac ggc ctg gcc 1597

Arg Arg Val Leu Val Asn Ala Ser Leu Gly Trp Pro Asn Gly Leu Ala

495 500 505

ctg gac ctg cag gag ggg aag ctc tac tgg gga gac gcc aag aca gac 1645

Leu Asp Leu Gln Glu Gly Lys Leu Tyr Trp Gly Asp Ala Lys Thr Asp

510 515 520 525

aag atc gag gtg atc aat gtt gat ggg acg aag agg cgg acc ctc ctg 1693

Lys Ile Glu Val Ile Asn Val Asp Gly Thr Lys Arg Arg Thr Leu Leu

530 535 540

gag gac aag ctc ccg cac att ttc ggg ttc acg ctg ctg ggg gac ttc 1741

Glu Asp Lys Leu Pro His Ile Phe Gly Phe Thr Leu Leu Gly Asp Phe

545 550 555

atc tac tgg act gac tgg cag cgc cgc agc atc gag cgg gtg cac aag 1789

Ile Tyr Trp Thr Asp Trp Gln Arg Arg Ser Ile Glu Arg Val His Lys

560 565 570

gtc aag gcc agc cgg gac gtc atc att gac cag ctg ccc gac ctg atg 1837

Val Lys Ala Ser Arg Asp Val Ile Ile Asp Gln Leu Pro Asp Leu Met

575 580 585

ggg ctc aaa gct gtg aat gtg gcc aag gtc gtc gga acc aac ccg tgt 1885

Gly Leu Lys Ala Val Asn Val Ala Lys Val Val Gly Thr Asn Pro Cys

590 595 600 605

gcg gac agg aac ggg ggg tgc agc cac ctg tgc ttc ttc aca ccc cac 1933

Ala Asp Arg Asn Gly Gly Cys Ser His Leu Cys Phe Phe Thr Pro His

610 615 620

gca acc cgg tgt ggc tgc ccc atc ggc ctg gag ctg ctg agt gac atg 1981

Ala Thr Arg Cys Gly Cys Pro Ile Gly Leu Glu Leu Leu Ser Asp Met

625 630 635

aag acc tgc atc gtg cct gag gcc ttc ttg gtc ttc acc agc aga gcc 2029

Lys Thr Cys Ile Val Pro Glu Ala Phe Leu Val Phe Thr Ser Arg Ala

640 645 650
gcc atc cac agg atc tcc ctc gag acc aat aac aac gac gtg gcc atc 2077
Ala Ile His Arg Ile Ser Leu Glu Thr Asn Asn Asn Asp Val Ala Ile
655 660 665
ccg ctc acg ggc gtc aag gag gcc tca gcc ctg gac ttt gat gtg tcc 2125
Pro Leu Thr Gly Val Lys Glu Ala Ser Ala Leu Asp Phe Asp Val Ser
670 675 680 685
aac aac cac atc tac tgg aca gac gtc agc ctg aag acc atc agc cgc 2173
Asn Asn His Ile Tyr Trp Thr Asp Val Ser Leu Lys Thr Ile Ser Arg
690 695 700
gcc ttc atg aac ggg agc tcg gtg gag cac gtg gtg gag ttt ggc ctt 2221
Ala Phe Met Asn Gly Ser Ser Val Glu His Val Val Glu Phe Gly Leu
705 710 715
gac tac ccc gag ggc atg gcc gtt gac tgg atg ggc aag aac ctc tac 2269
Asp Tyr Pro Glu Gly Met Ala Val Asp Trp Met Gly Lys Asn Leu Tyr
720 725 730
tgg gcc gac act ggg acc aac aga atc gaa gtg gcg cgg ctg gac ggg 2317
Trp Ala Asp Thr Gly Thr Asn Arg Ile Glu Val Ala Arg Leu Asp Gly
735 740 745
cag ttc cgg caa gtc ctc gtg tgg agg gac ttg gac aac ccg agg tcg 2365
Gln Phe Arg Gln Val Leu Val Trp Arg Asp Leu Asp Asn Pro Arg Ser
750 755 760 765
ctg gcc ctg gat ccc acc aag ggc tac atc tac tgg acc gag tgg ggc 2413

Leu Ala Leu Asp Pro Thr Lys Gly Tyr Ile Tyr Trp Thr Glu Trp Gly

770

775

780

ggc aag ccg agg atc gtg cgg gcc ttc atg gac ggg acc aac tgc atg 2461

Gly Lys Pro Arg Ile Val Arg Ala Phe Met Asp Gly Thr Asn Cys Met

785

790

795

acg ctg gtg gac aag gtg ggc cgg gcc aac gac ctc acc att gac tac 2509

Thr Leu Val Asp Lys Val Gly Arg Ala Asn Asp Leu Thr Ile Asp Tyr

800

805

810

gct gac cag cgc ctc tac tgg acc gac ctg gac acc aac atg atc gag 2557

Ala Asp Gln Arg Leu Tyr Trp Thr Asp Leu Asp Thr Asn Met Ile Glu

815

820

825

tcg tcc aac atg ctg ggt cag gag cgg gtc gtg att gcc gac gat ctc 2605

Ser Ser Asn Met Leu Gly Gln Glu Arg Val Val Ile Ala Asp Asp Leu

830

835

840

845

ccg cac ccg ttc ggt ctg acg cag tac agc gat tat atc tac tgg aca 2653

Pro His Pro Phe Gly Leu Thr Gln Tyr Ser Asp Tyr Ile Tyr Trp Thr

850

855

860

gac tgg aat ctg cac agc att gag cgg gcc gac aag act agc ggc cgg 2701

Asp Trp Asn Leu His Ser Ile Glu Arg Ala Asp Lys Thr Ser Gly Arg

865

870

875

aac cgc acc ctc atc cag ggc cac ctg gac ttc gtg atg gac atc ctg 2749

Asn Arg Thr Leu Ile Gln Gly His Leu Asp Phe Val Met Asp Ile Leu

880

885

890

gtg ttc cac tcc tcc cgc cag gat ggc ctc aat gac tgt atg cac aac 2797

Val Phe His Ser Ser Arg Gln Asp Gly Leu Asn Asp Cys Met His Asn

895 900 905

aac ggg cag tgt ggg cag ctg tgc ctt gcc atc ccc ggc ggc cac cgc 2845

Asn Gly Gln Cys Gly Gln Leu Cys Leu Ala Ile Pro Gly Gly His Arg

910 915 920 925

tgc ggc tgc gcc tca cac tac acc ctg gac ccc agc agc cgc aac tgc 2893

Cys Gly Cys Ala Ser His Tyr Thr Leu Asp Pro Ser Ser Arg Asn Cys

930 935 940

agc ccg ccc acc acc ttc ttg ctg ttc agc cag aaa tct gcc atc agt 2941

Ser Pro Pro Thr Thr Phe Leu Leu Phe Ser Gln Lys Ser Ala Ile Ser

945 950 955

cgg atg atc ccg gac gac cag cac agc ccg gat ctc atc ctg ccc ctg 2989

Arg Met Ile Pro Asp Asp Gln His Ser Pro Asp Leu Ile Leu Pro Leu

960 965 970

cat gga ctg agg aac gtc aaa gcc atc gac tat gac cca ctg gac aag 3037

His Gly Leu Arg Asn Val Lys Ala Ile Asp Tyr Asp Pro Leu Asp Lys

975 980 985

ttc atc tac tgg gtg gat ggg cgc cag aac atc aag cga gcc aag gac 3085

Phe Ile Tyr Trp Val Asp Gly Arg Gln Asn Ile Lys Arg Ala Lys Asp

990 995 1000 1005

gac ggg acc cag ccc ttt gtt ttg acc tct ctg agc caa ggc caa aac 3133

Asp Gly Thr Gln Pro Phe Val Leu Thr Ser Leu Ser Gln Gly Gln Asn

1010 1015 1020
cca gac agg cag ccc cac gac ctc agc atc gac atc tac agc cgg aca 3181
Pro Asp Arg Gln Pro His Asp Leu Ser Ile Asp Ile Tyr Ser Arg Thr
1025 1030 1035
ctg ttc tgg acg tgc gag gcc acc aat acc atc aac gtc cac agg ctg 3229
Leu Phe Trp Thr Cys Glu Ala Thr Asn Thr Ile Asn Val His Arg Leu
1040 1045 1050
agc ggg gaa gcc atg ggg gtg gtg ctg cgt ggg gac cgc gac aag ccc 3277
Ser Gly Glu Ala Met Gly Val Val Leu Arg Gly Asp Arg Asp Lys Pro
1055 1060 1065
agg gcc atc gtc gtc aac gcg gag cga ggg tac ctg tac ttc acc aac 3325
Arg Ala Ile Val Val Asn Ala Glu Arg Gly Tyr Leu Tyr Phe Thr Asn
1070 1075 1080 1085
atg cag gac cgg gca gcc aag atc gaa cgc gca gcc ctg gac ggc acc 3373
Met Gln Asp Arg Ala Ala Lys Ile Glu Arg Ala Ala Leu Asp Gly Thr
1090 1095 1100
gag cgc gag gtc ctc ttc acc acc ggc ctc atc cgc cct gtg gcc ctg 3421
Glu Arg Glu Val Leu Phe Thr Thr Gly Leu Ile Arg Pro Val Ala Leu
1105 1110 1115
gtg gtg gac aac aca ctg ggc aag ctg ttc tgg gtg gac gcg gac ctg 3469
Val Val Asp Asn Thr Leu Gly Lys Leu Phe Trp Val Asp Ala Asp Leu
1120 1125 1130
aag cgc att gag agc tgt gac ctg tca ggg gcc aac cgc ctg acc ctg 3517

Lys Arg Ile Glu Ser Cys Asp Leu Ser Gly Ala Asn Arg Leu Thr Leu

1135 1140 1145

gag gac gcc aac atc gtg cag cct ctg ggc ctg acc atc ctt ggc aag 3565

Glu Asp Ala Asn Ile Val Gln Pro Leu Gly Leu Thr Ile Leu Gly Lys

1150 1155 1160 1165

cat ctc tac tgg atc gac cgc cag cag cag atg atc gag cgt gtg gag 3613

His Leu Tyr Trp Ile Asp Arg Gln Gln Gln Met Ile Glu Arg Val Glu

1170 1175 1180

aag acc acc ggg gac aag cgg act cgc atc cag ggc cgt gtc gcc cac 3661

Lys Thr Thr Gly Asp Lys Arg Thr Arg Ile Gln Gly Arg Val Ala His

1185 1190 1195

ctc act ggc atc cat gca gtg gag gaa gtc agc ctg gag gag ttc tca 3709

Leu Thr Gly Ile His Ala Val Glu Glu Val Ser Leu Glu Glu Phe Ser

1200 1205 1210

gcc cac cca tgt gcc cgt gac aat ggt ggc tgc tcc cac atc tgt att 3757

Ala His Pro Cys Ala Arg Asp Asn Gly Gly Cys Ser His Ile Cys Ile

1215 1220 1225

gcc aag ggt gat ggg aca cca cgg tgc tca tgc cca gtc cac ctc gtg 3805

Ala Lys Gly Asp Gly Thr Pro Arg Cys Ser Cys Pro Val His Leu Val

1230 1235 1240 1245

ctc ctg cag aac ctg ctg acc tgt gga gag ccg ccc acc tgc tcc ccg 3853

Leu Leu Gln Asn Leu Leu Thr Cys Gly Glu Pro Pro Thr Cys Ser Pro

1250 1255 1260

gac cag ttt gca tgt gcc aca ggg gag atc gac tgt atc ccc ggg gcc 3901
Asp Gln Phe Ala Cys Ala Thr Gly Glu Ile Asp Cys Ile Pro Gly Ala
1265 1270 1275
tgg cgc tgt gac ggc ttt ccc gag tgc gat gac cag agc gac gag gag 3949
Trp Arg Cys Asp Gly Phe Pro Glu Cys Asp Asp Gln Ser Asp Glu Glu
1280 1285 1290
ggc tgc ccc gtg tgc tcc gcc gcc cag ttc ccc tgc gcg cgg ggt cag 3997
Gly Cys Pro Val Cys Ser Ala Ala Gln Phe Pro Cys Ala Arg Gly Gln
1295 1300 1305
tgt gtg gac ctg cgc ctg cgc tgc gac ggc gag gca gac tgt cag gac 4045
Cys Val Asp Leu Arg Leu Arg Cys Asp Gly Glu Ala Asp Cys Gln Asp
1310 1315 1320 1325
cgc tca gac gag gtg gac tgt gac gcc atc tgc ctg ccc aac cag ttc 4093
Arg Ser Asp Glu Val Asp Cys Asp Ala Ile Cys Leu Pro Asn Gln Phe
1330 1335 1340
cgg tgt gcg agc ggc cag tgt gtc ctc atc aaa cag cag tgc gac tcc 4141
Arg Cys Ala Ser Gly Gln Cys Val Leu Ile Lys Gln Gln Cys Asp Ser
1345 1350 1355
ttc ccc gac tgt atc gac ggc tcc gac gag ctc atg tgt gaa atc acc 4189
Phe Pro Asp Cys Ile Asp Gly Ser Asp Glu Leu Met Cys Glu Ile Thr
1360 1365 1370
aag ccg ccc tca gac gac agc ccg gcc cac agc agt gcc atc ggg ccc 4237
Lys Pro Pro Ser Asp Asp Ser Pro Ala His Ser Ser Ala Ile Gly Pro

1375 1380 1385
gtc att ggc atc atc ctc tct ctc ttc gtc atg ggt ggt gtc tat ttt 4285
Val Ile Gly Ile Ile Leu Ser Leu Phe Val Met Gly Gly Val Tyr Phe
1390 1395 1400 1405
gtg tgc cag cgc gtg gtg tgc cag cgc tat gcg ggg gcc aac ggg ccc 4333
Val Cys Gln Arg Val Val Cys Gln Arg Tyr Ala Gly Ala Asn Gly Pro
1410 1415 1420
ttc ccg cac gag tat gtc agc ggg acc ccg cac gtg ccc ctc aat ttc 4381
Phe Pro His Glu Tyr Val Ser Gly Thr Pro His Val Pro Leu Asn Phe
1425 1430 1435
ata gcc ccg ggc ggt tcc cag cat ggc ccc ttc aca ggc atc gca tgc 4429
Ile Ala Pro Gly Gly Ser Gln His Gly Pro Phe Thr Gly Ile Ala Cys
1440 1445 1450
gga aag tcc atg atg agc tcc gtg agc ctg atg ggg ggc cgg ggc ggg 4477
Gly Lys Ser Met Met Ser Ser Val Ser Leu Met Gly Gly Arg Gly Gly
1455 1460 1465
gtg ccc ctc tac gac cgg aac cac gtc aca ggg gcc tcg tcc agc agc 4525
Val Pro Leu Tyr Asp Arg Asn His Val Thr Gly Ala Ser Ser Ser Ser
1470 1475 1480 1485
tcg tcc agc acg aag gcc acg ctg tac ccg ccg atc ctg aac ccg ccg 4573
Ser Ser Ser Thr Lys Ala Thr Leu Tyr Pro Pro Ile Leu Asn Pro Pro
1490 1495 1500
ccc tcc ccg gcc acg gac ccc tcc ctg tac aac atg gac atg ttc tac 4621

Pro Ser Pro Ala Thr Asp Pro Ser Leu Tyr Asn Met Asp Met Phe Tyr

1505

1510

1515

tct tca aac att ccg gcc act gcg aga ccg tac agg ccc tac atc att 4669

Ser Ser Asn Ile Pro Ala Thr Ala Arg Pro Tyr Arg Pro Tyr Ile Ile

1520

1525

1530

cga gga atg gcg ccc ccg acg acg ccc tgc agc acc gac gtg tgt gac 4717

Arg Gly Met Ala Pro Pro Thr Thr Pro Cys Ser Thr Asp Val Cys Asp

1535

1540

1545

agc gac tac agc gcc agc cgc tgg aag gcc agc aag tac tac ctg gat 4765

Ser Asp Tyr Ser Ala Ser Arg Trp Lys Ala Ser Lys Tyr Tyr Leu Asp

1550

1555

1560

1565

ttg aac tcg gac tca gac ccc tat cca ccc cca ccc acg ccc cac agc 4813

Leu Asn Ser Asp Ser Asp Pro Tyr Pro Pro Pro Pro Thr Pro His Ser

1570

1575

1580

cag tac ctg tcg gcg gag gac agc tgc ccg ccc tcg ccc gcc acc gag 4861

Gln Tyr Leu Ser Ala Glu Asp Ser Cys Pro Pro Ser Pro Ala Thr Glu

1585

1590

1595

agg agc tac ttc cat ctc ttc ccg ccc cct ccg tcc ccc tgc acg gac 4909

Arg Ser Tyr Phe His Leu Phe Pro Pro Pro Pro Ser Pro Cys Thr Asp

1600

1605

1610

tca tcc tgacctcggc cgggccactc tggcttctct gtgccctgt aaatagtttt 4965

Ser Ser

1615

aaatatgaac aaagaaaaa atatatttta tgatttaaaa aataaatata attgggatt 5025
taaaaacatg agaaatgtga actgtgatgg ggtgggcagg gctgggagaa cttgtacag 5085
tggagaaata ttataaact taattttgta aaaca 5120

<210> 3

<211> 1615

<212> PRT

<213> Homo sapiens

<400> 3

Met Glu Ala Ala Pro Pro Gly Pro Pro Trp Pro Leu Leu Leu Leu

1 5 10 15

Leu Leu Leu Leu Ala Leu Cys Gly Cys Pro Ala Pro Ala Ala Ser

20 25 30

Pro Leu Leu Leu Phe Ala Asn Arg Arg Asp Val Arg Leu Val Asp Ala

35 40 45

Gly Gly Val Lys Leu Glu Ser Thr Ile Val Val Ser Gly Leu Glu Asp

50 55 60

Ala Ala Ala Val Asp Phe Gln Phe Ser Lys Gly Ala Val Tyr Trp Thr

65 70 75 80

Asp Val Ser Glu Glu Ala Ile Lys Gln Thr Tyr Leu Asn Gln Thr Gly

85 90 95

Ala Ala Val Gln Asn Val Val Ile Ser Gly Leu Val Ser Pro Asp Gly

100 105 110
Leu Ala Cys Asp Trp Val Gly Lys Lys Leu Tyr Trp Thr Asp Ser Glu
115 120 125
Thr Asn Arg Ile Glu Val Ala Asn Leu Asn Gly Thr Ser Arg Lys Val
130 135 140
Leu Phe Trp Gln Asp Leu Asp Gln Pro Lys Ala Ile Ala Leu Asp Pro
145 150 155 160
Ala His Gly Tyr Met Tyr Trp Thr Asp Trp Gly Glu Thr Pro Arg Ile
165 170 175
Glu Arg Ala Gly Met Asp Gly Ser Thr Arg Lys Ile Ile Val Asp Ser
180 185 190
Asp Ile Tyr Trp Pro Asn Gly Leu Thr Ile Asp Leu Glu Glu Gln Lys
195 200 205
Leu Tyr Trp Ala Asp Ala Lys Leu Ser Phe Ile His Arg Ala Asn Leu
210 215 220
Asp Gly Ser Phe Arg Gln Lys Val Val Glu Gly Ser Leu Thr His Pro
225 230 235 240
Phe Ala Leu Thr Leu Ser Gly Asp Thr Leu Tyr Trp Thr Asp Trp Gln
245 250 255
Thr Arg Ser Ile His Ala Cys Asn Lys Arg Thr Gly Gly Lys Arg Lys
260 265 270
Glu Ile Leu Ser Ala Leu Tyr Ser Pro Met Asp Ile Gln Val Leu Ser
275 280 285

Gln Glu Arg Gln Pro Phe Phe His Thr Arg Cys Glu Glu Asp Asn Gly

290 295 300

Gly Trp Ser His Leu Cys Leu Leu Ser Pro Ser Glu Pro Phe Tyr Thr

305 310 315 320

Cys Ala Cys Pro Thr Gly Val Gln Met Gln Asp Asn Gly Arg Thr Cys

325 330 335

Lys Ala Gly Ala Glu Glu Val Leu Leu Leu Ala Arg Arg Thr Asp Leu

340 345 350

Arg Arg Ile Ser Leu Asp Thr Pro Asp Phe Thr Asp Ile Val Leu Gln

355 360 365

Val Asp Asp Ile Arg His Ala Ile Ala Ile Asp Tyr Asp Pro Leu Glu

370 375 380

Gly Tyr Val Tyr Trp Thr Asp Asp Glu Val Arg Ala Ile Arg Arg Ala

385 390 395 400

Tyr Leu Asp Gly Ser Gly Ala Gln Thr Leu Val Asn Thr Glu Ile Asn

405 410 415

Asp Pro Asp Gly Ile Ala Val Asp Trp Val Ala Arg Asn Leu Tyr Trp

420 425 430

Thr Asp Thr Gly Thr Asp Arg Ile Glu Val Thr Arg Leu Asn Gly Thr

435 440 445

Ser Arg Lys Ile Leu Val Ser Glu Asp Leu Asp Glu Pro Arg Ala Ile

450 455 460

Ala Leu His Pro Val Met Gly Leu Met Tyr Trp Thr Asp Trp Gly Glu

465 470 475 480
Asn Pro Lys Ile Glu Cys Ala Asn Leu Asp Gly Gln Glu Arg Arg Val

 485 490 495
Leu Val Asn Ala Ser Leu Gly Trp Pro Asn Gly Leu Ala Leu Asp Leu

 500 505 510
Gln Glu Gly Lys Leu Tyr Trp Gly Asp Ala Lys Thr Asp Lys Ile Glu

 515 520 525
Val Ile Asn Val Asp Gly Thr Lys Arg Arg Thr Leu Leu Glu Asp Lys

 530 535 540
Leu Pro His Ile Phe Gly Phe Thr Leu Leu Gly Asp Phe Ile Tyr Trp

545 550 555 560
Thr Asp Trp Gln Arg Arg Ser Ile Glu Arg Val His Lys Val Lys Ala

 565 570 575
Ser Arg Asp Val Ile Ile Asp Gln Leu Pro Asp Leu Met Gly Leu Lys

 580 585 590
Ala Val Asn Val Ala Lys Val Val Gly Thr Asn Pro Cys Ala Asp Arg

 595 600 605
Asn Gly Gly Cys Ser His Leu Cys Phe Phe Thr Pro His Ala Thr Arg

 610 615 620
Cys Gly Cys Pro Ile Gly Leu Glu Leu Leu Ser Asp Met Lys Thr Cys

625 630 635 640
Ile Val Pro Glu Ala Phe Leu Val Phe Thr Ser Arg Ala Ala Ile His

 645 650 655

Arg Ile Ser Leu Glu Thr Asn Asn Asn Asp Val Ala Ile Pro Leu Thr

660 665 670

Gly Val Lys Glu Ala Ser Ala Leu Asp Phe Asp Val Ser Asn Asn His

675 680 685

Ile Tyr Trp Thr Asp Val Ser Leu Lys Asn Ile Ser Arg Ala Phe Met

690 695 700

Asn Gly Ser Ser Val Glu His Val Val Glu Phe Gly Leu Asp Tyr Pro

705 710 715 720

Glu Gly Met Ala Val Asp Trp Met Gly Lys Asn Leu Tyr Trp Ala Asp

725 730 735

Thr Gly Thr Asn Arg Ile Glu Val Ala Arg Leu Asp Gly Gln Phe Arg

740 745 750

Gln Val Leu Val Trp Arg Asp Leu Asp Asn Pro Arg Ser Leu Ala Leu

755 760 765

Asp Pro Thr Lys Gly Tyr Ile Tyr Trp Thr Glu Trp Gly Gly Lys Pro

770 775 780

Arg Ile Val Arg Ala Phe Met Asp Gly Thr Asn Cys Met Thr Leu Val

785 790 795 800

Asp Lys Val Gly Arg Ala Asn Asp Leu Thr Ile Asp Tyr Ala Asp Gln

805 810 815

Arg Leu Tyr Trp Thr Asp Leu Asp Thr Asn Met Ile Glu Ser Ser Asn

820 825 830

Met Leu Gly Gln Glu Arg Val Val Ile Ala Asp Asp Leu Pro His Pro

835 840 845
Phe Gly Leu Thr Gln Tyr Ser Asp Tyr Ile Tyr Trp Thr Asp Trp Asn
850 855 860
Leu His Ser Ile Glu Arg Ala Asp Lys Thr Ser Gly Arg Asn Arg Thr
865 870 875 880
Leu Ile Gln Gly His Leu Asp Phe Val Met Asp Ile Leu Val Phe His
885 890 895
Ser Ser Arg Gln Asp Gly Leu Asn Asp Cys Met His Asn Asn Gly Gln
900 905 910
Cys Gly Gln Leu Cys Leu Ala Ile Pro Gly Gly His Arg Cys Gly Cys
915 920 925
Ala Ser His Tyr Thr Leu Asp Pro Ser Ser Arg Asn Cys Ser Pro Pro
930 935 940
Thr Thr Phe Leu Leu Phe Ser Gln Lys Ser Ala Ile Ser Arg Met Ile
945 950 955 960
Pro Asp Asp Gln His Ser Pro Asp Leu Ile Leu Pro Leu His Gly Leu
965 970 975
Arg Asn Val Lys Ala Ile Asp Tyr Asp Pro Leu Asp Lys Phe Ile Tyr
980 985 990
Trp Val Asp Gly Arg Gln Asn Ile Lys Arg Ala Lys Asp Asp Gly Thr
995 1000 1005
Gln Pro Phe Val Leu Thr Ser Leu Ser Gln Gly Gln Asn Pro Asp Arg
1010 1015 1020

Gln Pro His Asp Leu Ser Ile Asp Ile Tyr Ser Arg Thr Leu Phe Trp

1025 1030 1035 1040

Thr Cys Glu Ala Thr Asn Thr Ile Asn Val His Arg Leu Ser Gly Glu

1045 1050 1055

Ala Met Gly Val Val Leu Arg Gly Asp Arg Asp Lys Pro Arg Ala Ile

1060 1065 1070

Val Val Asn Ala Glu Arg Gly Tyr Leu Tyr Phe Thr Asn Met Gln Asp

1075 1080 1085

Arg Ala Ala Lys Ile Glu Arg Ala Ala Leu Asp Gly Thr Glu Arg Glu

1090 1095 1100

Val Leu Phe Thr Thr Gly Leu Ile Arg Pro Val Ala Leu Val Val Asp

1105 1110 1115 1120

Asn Thr Leu Gly Lys Leu Phe Trp Val Asp Ala Asp Leu Lys Arg Ile

1125 1130 1135

Glu Ser Cys Asp Leu Ser Gly Ala Asn Arg Leu Thr Leu Glu Asp Ala

1140 1145 1150

Asn Ile Val Gln Pro Leu Gly Leu Thr Ile Leu Gly Lys His Leu Tyr

1155 1160 1165

Trp Ile Asp Arg Gln Gln Gln Met Ile Glu Arg Val Glu Lys Thr Thr

1170 1175 1180

Gly Asp Lys Arg Thr Arg Ile Gln Gly Arg Val Ala His Leu Thr Gly

1185 1190 1195 1200

Ile His Ala Val Glu Glu Val Ser Leu Glu Glu Phe Ser Ala His Pro

1205 1210 1215
Cys Ala Arg Asp Asn Gly Gly Cys Ser His Ile Cys Ile Ala Lys Gly
1220 1225 1230
Asp Gly Thr Pro Arg Cys Ser Cys Pro Val His Leu Val Leu Leu Gln
1235 1240 1245
Asn Leu Leu Thr Cys Gly Glu Pro Pro Thr Cys Ser Pro Asp Gln Phe
1250 1255 1260
Ala Cys Ala Thr Gly Glu Ile Asp Cys Ile Pro Gly Ala Trp Arg Cys
1265 1270 1275 1280
Asp Gly Phe Pro Glu Cys Asp Asp Gln Ser Asp Glu Glu Gly Cys Pro
1285 1290 1295
Val Cys Ser Ala Ala Gln Phe Pro Cys Ala Arg Gly Gln Cys Val Asp
1300 1305 1310
Leu Arg Leu Arg Cys Asp Gly Glu Ala Asp Cys Gln Asp Arg Ser Asp
1315 1320 1325
Glu Val Asp Cys Asp Ala Ile Cys Leu Pro Asn Gln Phe Arg Cys Ala
1330 1335 1340
Ser Gly Gln Cys Val Leu Ile Lys Gln Gln Cys Asp Ser Phe Pro Asp
1345 1350 1355 1360
Cys Ile Asp Gly Ser Asp Glu Leu Met Cys Glu Ile Thr Lys Pro Pro
1365 1370 1375
Ser Asp Asp Ser Pro Ala His Ser Ser Ala Ile Gly Pro Val Ile Gly
1380 1385 1390

Ile Ile Leu Ser Leu Phe Val Met Gly Gly Val Tyr Phe Val Cys Gln

1395 1400 1405

Arg Val Val Cys Gln Arg Tyr Ala Gly Ala Asn Gly Pro Phe Pro His

1410 1415 1420

Glu Tyr Val Ser Gly Thr Pro His Val Pro Leu Asn Phe Ile Ala Pro

1425 1430 1435 1440

Gly Gly Ser Gln His Gly Pro Phe Thr Gly Ile Ala Cys Gly Lys Ser

1445 1450 1455

Met Met Ser Ser Val Ser Leu Met Gly Gly Arg Gly Gly Val Pro Leu

1460 1465 1470

Tyr Asp Arg Asn His Val Thr Gly Ala Ser Ser Ser Ser Ser Ser

1475 1480 1485

Thr Lys Ala Thr Leu Tyr Pro Pro Ile Leu Asn Pro Pro Pro Ser Pro

1490 1495 1500

Ala Thr Asp Pro Ser Leu Tyr Asn Met Asp Met Phe Tyr Ser Ser Asn

1505 1510 1515 1520

Ile Pro Ala Thr Ala Arg Pro Tyr Arg Pro Tyr Ile Ile Arg Gly Met

1525 1530 1535

Ala Pro Pro Thr Thr Pro Cys Ser Thr Asp Val Cys Asp Ser Asp Tyr

1540 1545 1550

Ser Ala Ser Arg Trp Lys Ala Ser Lys Tyr Tyr Leu Asp Leu Asn Ser

1555 1560 1565

Asp Ser Asp Pro Tyr Pro Pro Pro Pro Thr Pro His Ser Gln Tyr Leu

1570 1575 1580
Ser Ala Glu Asp Ser Cys Pro Pro Ser Pro Ala Thr Glu Arg Ser Tyr
1585 1590 1595 1600
Phe His Leu Phe Pro Pro Pro Pro Ser Pro Cys Thr Asp Ser Ser
1605 1610 1615

<210> 4

<211> 1615

<212> PRT

<213> Homo sapiens

<400> 4

Met Glu Ala Ala Pro Pro Gly Pro Pro Trp Pro Leu Leu Leu Leu Leu
1 5 10 15
Leu Leu Leu Leu Ala Leu Cys Gly Cys Pro Ala Pro Ala Ala Ala Ser
20 25 30
Pro Leu Leu Leu Phe Ala Asn Arg Arg Asp Val Arg Leu Val Asp Ala
35 40 45
Gly Gly Val Lys Leu Glu Ser Thr Ile Val Val Ser Gly Leu Glu Asp
50 55 60
Ala Ala Ala Val Asp Phe Gln Phe Ser Lys Gly Ala Val Tyr Trp Thr
65 70 75 80

Asp Val Ser Glu Glu Ala Ile Lys Gln Thr Tyr Leu Asn Gln Thr Gly

85 90 95

Ala Ala Val Gln Asn Val Val Ile Ser Gly Leu Val Ser Pro Asp Gly

100 105 110

Leu Ala Cys Asp Trp Val Gly Lys Lys Leu Tyr Trp Thr Asp Ser Glu

115 120 125

Thr Asn Arg Ile Glu Val Ala Asn Leu Asn Gly Thr Ser Arg Lys Val

130 135 140

Leu Phe Trp Gln Asp Leu Asp Gln Pro Lys Ala Ile Ala Leu Asp Pro

145 150 155 160

Ala His Gly Tyr Met Tyr Trp Thr Asp Trp Val Glu Thr Pro Arg Ile

165 170 175

Glu Arg Ala Gly Met Asp Gly Ser Thr Arg Lys Ile Ile Val Asp Ser

180 185 190

Asp Ile Tyr Trp Pro Asn Gly Leu Thr Ile Asp Leu Glu Glu Gln Lys

195 200 205

Leu Tyr Trp Ala Asp Ala Lys Leu Ser Phe Ile His Arg Ala Asn Leu

210 215 220

Asp Gly Ser Phe Arg Gln Lys Val Val Glu Gly Ser Leu Thr His Pro

225 230 235 240

Phe Ala Leu Thr Leu Ser Gly Asp Thr Leu Tyr Trp Thr Asp Trp Gln

245 250 255

Thr Arg Ser Ile His Ala Cys Asn Lys Arg Thr Gly Gly Lys Arg Lys

260 265 270
Glu Ile Leu Ser Ala Leu Tyr Ser Pro Met Asp Ile Gln Val Leu Ser

275 280 285
Gln Glu Arg Gln Pro Phe Phe His Thr Arg Cys Glu Glu Asp Asn Gly

290 295 300
Gly Trp Ser His Leu Cys Leu Leu Ser Pro Ser Glu Pro Phe Tyr Thr

305 310 315 320
Cys Ala Cys Pro Thr Gly Val Gln Met Gln Asp Asn Gly Arg Thr Cys

325 330 335
Lys Ala Gly Ala Glu Glu Val Leu Leu Leu Ala Arg Arg Thr Asp Leu

340 345 350
Arg Arg Ile Ser Leu Asp Thr Pro Asp Phe Thr Asp Ile Val Leu Gln

355 360 365
Val Asp Asp Ile Arg His Ala Ile Ala Ile Asp Tyr Asp Pro Leu Glu

370 375 380
Gly Tyr Val Tyr Trp Thr Asp Asp Glu Val Arg Ala Ile Arg Arg Ala

385 390 395 400
Tyr Leu Asp Gly Ser Gly Ala Gln Thr Leu Val Asn Thr Glu Ile Asn

405 410 415
Asp Pro Asp Gly Ile Ala Val Asp Trp Val Ala Arg Asn Leu Tyr Trp

420 425 430
Thr Asp Thr Gly Thr Asp Arg Ile Glu Val Thr Arg Leu Asn Gly Thr

435 440 445

Ser Arg Lys Ile Leu Val Ser Glu Asp Leu Asp Glu Pro Arg Ala Ile

450 455 460

Ala Leu His Pro Val Met Gly Leu Met Tyr Trp Thr Asp Trp Gly Glu

465 470 475 480

Asn Pro Lys Ile Glu Cys Ala Asn Leu Asp Gly Gln Glu Arg Arg Val

485 490 495

Leu Val Asn Ala Ser Leu Gly Trp Pro Asn Gly Leu Ala Leu Asp Leu

500 505 510

Gln Glu Gly Lys Leu Tyr Trp Gly Asp Ala Lys Thr Asp Lys Ile Glu

515 520 525

Val Ile Asn Val Asp Gly Thr Lys Arg Arg Thr Leu Leu Glu Asp Lys

530 535 540

Leu Pro His Ile Phe Gly Phe Thr Leu Leu Gly Asp Phe Ile Tyr Trp

545 550 555 560

Thr Asp Trp Gln Arg Arg Ser Ile Glu Arg Val His Lys Val Lys Ala

565 570 575

Ser Arg Asp Val Ile Ile Asp Gln Leu Pro Asp Leu Met Gly Leu Lys

580 585 590

Ala Val Asn Val Ala Lys Val Val Gly Thr Asn Pro Cys Ala Asp Arg

595 600 605

Asn Gly Gly Cys Ser His Leu Cys Phe Phe Thr Pro His Ala Thr Arg

610 615 620

Cys Gly Cys Pro Ile Gly Leu Glu Leu Leu Ser Asp Met Lys Thr Cys

625 630 635 640
Ile Val Pro Glu Ala Phe Leu Val Phe Thr Ser Arg Ala Ala Ile His

645 650 655
Arg Ile Ser Leu Glu Thr Asn Asn Asn Asp Val Ala Ile Pro Leu Thr

660 665 670
Gly Val Lys Glu Ala Ser Ala Leu Asp Phe Asp Val Ser Asn Asn His

675 680 685
Ile Tyr Trp Thr Asp Val Ser Leu Lys Asn Ile Ser Arg Ala Phe Met

690 695 700
Asn Gly Ser Ser Val Glu His Val Val Glu Phe Gly Leu Asp Tyr Pro

705 710 715 720
Glu Gly Met Ala Val Asp Trp Met Gly Lys Asn Leu Tyr Trp Ala Asp

725 730 735
Thr Gly Thr Asn Arg Ile Glu Val Ala Arg Leu Asp Gly Gln Phe Arg

740 745 750
Gln Val Leu Val Trp Arg Asp Leu Asp Asn Pro Arg Ser Leu Ala Leu

755 760 765
Asp Pro Thr Lys Gly Tyr Ile Tyr Trp Thr Glu Trp Gly Gly Lys Pro

770 775 780
Arg Ile Val Arg Ala Phe Met Asp Gly Thr Asn Cys Met Thr Leu Val

785 790 795 800
Asp Lys Val Gly Arg Ala Asn Asp Leu Thr Ile Asp Tyr Ala Asp Gln

805 810 815

Arg Leu Tyr Trp Thr Asp Leu Asp Thr Asn Met Ile Glu Ser Ser Asn

820 825 830

Met Leu Gly Gln Glu Arg Val Val Ile Ala Asp Asp Leu Pro His Pro

835 840 845

Phe Gly Leu Thr Gln Tyr Ser Asp Tyr Ile Tyr Trp Thr Asp Trp Asn

850 855 860

Leu His Ser Ile Glu Arg Ala Asp Lys Thr Ser Gly Arg Asn Arg Thr

865 870 875 880

Leu Ile Gln Gly His Leu Asp Phe Val Met Asp Ile Leu Val Phe His

885 890 895

Ser Ser Arg Gln Asp Gly Leu Asn Asp Cys Met His Asn Asn Gly Gln

900 905 910

Cys Gly Gln Leu Cys Leu Ala Ile Pro Gly Gly His Arg Cys Gly Cys

915 920 925

Ala Ser His Tyr Thr Leu Asp Pro Ser Ser Arg Asn Cys Ser Pro Pro

930 935 940

Thr Thr Phe Leu Leu Phe Ser Gln Lys Ser Ala Ile Ser Arg Met Ile

945 950 955 960

Pro Asp Asp Gln His Ser Pro Asp Leu Ile Leu Pro Leu His Gly Leu

965 970 975

Arg Asn Val Lys Ala Ile Asp Tyr Asp Pro Leu Asp Lys Phe Ile Tyr

980 985 990

Trp Val Asp Gly Arg Gln Asn Ile Lys Arg Ala Lys Asp Asp Gly Thr

995 1000 1005
Gln Pro Phe Val Leu Thr Ser Leu Ser Gln Gly Gln Asn Pro Asp Arg
1010 1015 1020
Gln Pro His Asp Leu Ser Ile Asp Ile Tyr Ser Arg Thr Leu Phe Trp
1025 1030 1035 1040
Thr Cys Glu Ala Thr Asn Thr Ile Asn Val His Arg Leu Ser Gly Glu
1045 1050 1055
Ala Met Gly Val Val Leu Arg Gly Asp Arg Asp Lys Pro Arg Ala Ile
1060 1065 1070
Val Val Asn Ala Glu Arg Gly Tyr Leu Tyr Phe Thr Asn Met Gln Asp
1075 1080 1085
Arg Ala Ala Lys Ile Glu Arg Ala Ala Leu Asp Gly Thr Glu Arg Glu
1090 1095 1100
Val Leu Phe Thr Thr Gly Leu Ile Arg Pro Val Ala Leu Val Val Asp
1105 1110 1115 1120
Asn Thr Leu Gly Lys Leu Phe Trp Val Asp Ala Asp Leu Lys Arg Ile
1125 1130 1135
Glu Ser Cys Asp Leu Ser Gly Ala Asn Arg Leu Thr Leu Glu Asp Ala
1140 1145 1150
Asn Ile Val Gln Pro Leu Gly Leu Thr Ile Leu Gly Lys His Leu Tyr
1155 1160 1165
Trp Ile Asp Arg Gln Gln Gln Met Ile Glu Arg Val Glu Lys Thr Thr
1170 1175 1180

Gly Asp Lys Arg Thr Arg Ile Gln Gly Arg Val Ala His Leu Thr Gly

1185 1190 1195 1200

Ile His Ala Val Glu Glu Val Ser Leu Glu Glu Phe Ser Ala His Pro

1205 1210 1215

Cys Ala Arg Asp Asn Gly Gly Cys Ser His Ile Cys Ile Ala Lys Gly

1220 1225 1230

Asp Gly Thr Pro Arg Cys Ser Cys Pro Val His Leu Val Leu Leu Gln

1235 1240 1245

Asn Leu Leu Thr Cys Gly Glu Pro Pro Thr Cys Ser Pro Asp Gln Phe

1250 1255 1260

Ala Cys Ala Thr Gly Glu Ile Asp Cys Ile Pro Gly Ala Trp Arg Cys

1265 1270 1275 1280

Asp Gly Phe Pro Glu Cys Asp Asp Gln Ser Asp Glu Glu Gly Cys Pro

1285 1290 1295

Val Cys Ser Ala Ala Gln Phe Pro Cys Ala Arg Gly Gln Cys Val Asp

1300 1305 1310

Leu Arg Leu Arg Cys Asp Gly Glu Ala Asp Cys Gln Asp Arg Ser Asp

1315 1320 1325

Glu Val Asp Cys Asp Ala Ile Cys Leu Pro Asn Gln Phe Arg Cys Ala

1330 1335 1340

Ser Gly Gln Cys Val Leu Ile Lys Gln Gln Cys Asp Ser Phe Pro Asp

1345 1350 1355 1360

Cys Ile Asp Gly Ser Asp Glu Leu Met Cys Glu Ile Thr Lys Pro Pro

1365 1370 1375
Ser Asp Asp Ser Pro Ala His Ser Ser Ala Ile Gly Pro Val Ile Gly
1380 1385 1390
Ile Ile Leu Ser Leu Phe Val Met Gly Gly Val Tyr Phe Val Cys Gln
1395 1400 1405
Arg Val Val Cys Gln Arg Tyr Ala Gly Ala Asn Gly Pro Phe Pro His
1410 1415 1420
Glu Tyr Val Ser Gly Thr Pro His Val Pro Leu Asn Phe Ile Ala Pro
1425 1430 1435 1440
Gly Gly Ser Gln His Gly Pro Phe Thr Gly Ile Ala Cys Gly Lys Ser
1445 1450 1455
Met Met Ser Ser Val Ser Leu Met Gly Gly Arg Gly Gly Val Pro Leu
1460 1465 1470
Tyr Asp Arg Asn His Val Thr Gly Ala Ser Ser Ser Ser Ser Ser Ser
1475 1480 1485
Thr Lys Ala Thr Leu Tyr Pro Pro Ile Leu Asn Pro Pro Pro Ser Pro
1490 1495 1500
Ala Thr Asp Pro Ser Leu Tyr Asn Met Asp Met Phe Tyr Ser Ser Asn
1505 1510 1515 1520
Ile Pro Ala Thr Ala Arg Pro Tyr Arg Pro Tyr Ile Ile Arg Gly Met
1525 1530 1535
Ala Pro Pro Thr Thr Pro Cys Ser Thr Asp Val Cys Asp Ser Asp Tyr
1540 1545 1550

Ser Ala Ser Arg Trp Lys Ala Ser Lys Tyr Tyr Leu Asp Leu Asn Ser

1555 1560 1565

Asp Ser Asp Pro Tyr Pro Pro Pro Pro Thr Pro His Ser Gln Tyr Leu

1570 1575 1580

Ser Ala Glu Asp Ser Cys Pro Pro Ser Pro Ala Thr Glu Arg Ser Tyr

1585 1590 1595 1600

Phe His Leu Phe Pro Pro Pro Pro Ser Pro Cys Thr Asp Ser Ser

1605 1610 1615

<210> 5

<211> 3096

<212> DNA

<213> Homo sapiens

<400> 5

catcttctca cacgatctct cgcttcgcac tccttccttt gattggtttt caccatttac 60

tcagacgacg gtcttcttc gatcttgca cattcttcta tcattacta cttcataacc 120

cagctccgc ccctaattt catgcgcgga tggcccatc cgtggtgaaa attcccttct 180

actctgctaa tctgctgttc tcttccctc ccgtcgggtt ctgctcctgc cacgttctcc 240

ccttcccca ccaaaggctg ggttttcttt gtcagggtc ctttccctt tggaagaagg 300

ggggctgtat ggccttggtg cgaggccctc cagtgcacagg atccccatc acccagagtt 360

ccacaggccc tggtagggag gagggggagc agaagaggag gtgccatctt tgcctgctgg 420

ggaagggcag gggccacca cacagagctc tcccatttgc tgtggaccct ggggccactg 480
cccagttcct tccaaaggaa agccagctcc ccaggtggtg ggagagtgat atggcttct 540
cttaaaactta gggaattgag tgtgtggttg ctctaaagt ccttagaagc cgggagcggc 600
tcttggaag agcctgcctg ccacagcggg ccttaccctg gctgtgcca cagatgtccc 660
tggggcctgc cgtcctgcc cggctctcct ggctccccc ggtgtgggtt gggaaaagca 720
cagcaaatta aaaaacacct ccatctctgg ctttgaaga atgcatctga acagccgaga 780
gtgtaaaccg tggtagaatg tggctttcc agttgggga gaagcagggc agagctgggg 840
ctttgtacc cagggttcc aagagctcct gcctccctcg gctgggctgg ccagggcccc 900
ccgctgggac ctccagctgt aatagggaag gttttactgg gttgctggcc actgtggact 960
gcccctaagg gcaggtatgc ctgcctttac ccgggttccc ctctgcctg gaagatacag 1020
cccatgggag gcctgttgc tgtgggatcc tccagcatca gagacactgg ggccagcgtc 1080
tgcctggtga ggtgcaggcc tggcaggccc ggtcccccac ctgcttgagc acccacggtg 1140
gtgggggctc gctgcctccc gagacaatct atgtcattgt tgtccaagga agctaattta 1200
gagtagaaag ticcgtgtcc agtcccactc tgtgcgtgtg ttagcagggg actctcgggc 1260
cggagctggg tccaccttg tagggggact tcatggggcc tgggcgacag cactgtgtat 1320
ttgtgtgtgt gtgtgttgt gtgtgtgtgt gtctgaggag gtggaccagt ttctaaaag 1380
gcctgtgacc ccaagaacca aggaattica gcctgggtgg atcacacctt cactggtgag 1440
tgggacaagc tgggggccct cgccacagga gcagccaggg catggggcac agttggcctc 1500
attcacaaaa tgggagtata agtgatccct gctctggcgg ccaggacgat gagtgggaac 1560
acaccgtgtg ggggctgcct ggctgggtg tgccgcgggt gtccttgtt gtgatggtc 1620
cacctgcttg tgccaccagt gccctctggg tctcacacac aactctctc ccagcgaagg 1680
cccctctgc cctcaggcct cagtgtgtgt tccgtctcgg aaggccccag gagctcctgc 1740
atcctgggcg tgattcctgt gtgcctgcag accccctcgc ggctgccatc tcctccttg 1800

gtgcacctgt tggccagacc tcctggtagc ggggtgctgca ctcccctgaa tgtgccgggg 1860
cctgggggca gggacctggg ctctccctc actgagtggg gggaaactcag tgtcttggag 1920
ttggggtgcc tgcaggctgg gtggtgcagg tgaaatgcag acctctcagc tgggtgtcca 1980
gagcagctgc ctccccgc ccgagggact tcaccgcag ccagtcagg ggtggcgct 2040
gggtgcatcg cccgcaggct gggtaggggt ggagcctggg tggcctgcc tgtgagctgc 2100
atagttgtcg ccttgaccc tgagtttct tcgtatctg ttggacctg ttggggcag 2160
gcaggggatg agatctgaag ataaatgcct tagctgtgac catctcttt tgtgagaggt 2220
caatgtccag ttccgtgca gttataacat cccattttt gatttcttt tatttttcc 2280
ttttcttt ttagatggag tctcgtctg tcaccaggc tggagtcaa tggggtgacc 2340
tcagctcact gcaacctcca cttctcgggt tcaagtatt ctctgcctc agcctctga 2400
ctagcagggg ttacaggcgt gagccaccac gccagctaa ttttgtatt ttagtagag 2460
gcaaggttc gtcattgtg ccaggctgggt ctcaaactcc tggccttaag tgatctgcc 2520
gcctcggcct cccaaagtgc tgagatgaca ggtgtgagcc accgtgcccg gccagaact 2580
cttaattcc cacctgaaac ttccgcctt aagcaggctc ccagctccc tcccctagtc 2640
cctggtccca ccattctgt tctgtctca atgaattgc ctaccgtaag tacctcatat 2700
aaattgaatc ataaagtatt tgtctttta tatctggctt atttactta gcataacatt 2760
cttaagttc atccatgtg tagcatgtg cagaatctct ctctttttt tttttttt 2820
tttttttt ttgtcagac agagtctgc tctgcatct agactggagt tcagtggcac 2880
gatctcgggt cactgcaaca tctgcctcct gggccaage aattctctg cctcagctc 2940
cttagcagct ggaactacag gcgcgtgcca ccatgcctg ctaattttg tattttatg 3000
tggaggcagg gtttaccat ctggccagg ctggtctga attctggtc ttaccacgg 3060
gggcccgaag gacccgggca aagcgtggag gggagg 3096

<210> 6

<211> 26928

<212> DNA

<213> Homo sapiens

<220>

<221> unsure

<222> (12044),(12489),(26433),(26434),(26435),(26436),(26439),(26441)

<223> Identity of nucleotide sequences at the above locations are unknown.

<400> 6

gaagaccaag ggcacacagc gaggcagttt cagggcgggc agcctggggc cccacggggc 60
ggccccggac acttgttctc acctgtggag ggcagagaag ggaacaggga gagaagtggc 120
cggctgggag tggaggtggg ttgaggttt tactgtaaac taaatgtgta ccctctacct 180
tagttatgaa ttatgagaca cgaagactgc gaaacagaca cactcctcta aaagtgcctc 240
taggctgaca gggagaaagt cccgccaggc tcccagacgc cacctttgag tccttcaaca 300
agcccgccag ggcctcttgc ccaccggtgt cagctcagcc actgaaccct ccaggaagaa 360
gacgtgctgg taggagaaga atctcaccca ggcacagcct ggaaggggca cagaaggggc 420
tccggaacca gcaagcccaa gtggaactc ccagtctgct actttctaga acgactgtgc 480
ccttggcggg tctaagtaga acctctccgc gcactcttc ctctttgta aagtggggac 540
agcaatggcc accttgcagg ttcagagagg gcttgcagta cctcacagaa ctgagtcccc 600
gtgaacgtgt gtgttctcc agatttga cagctttgcc aggctggagt caggctgaac 660
gcctctgccc tcatggggtt tatattctag gaagaccaac aaaaacaaga agacggaaaa 720

ttaaaacaac aaaagcccca ttgacaggcc gtgaagaatg ccatgaaaa tgaatggcgt 780
tgtgctgcag tctttggga aacgggctta cggaaagaag gacactgag ctgctaccaa 840
tgagcagccg tccggtggga gggcagtca ggaagagcag acatccactg aggaggcgct 900
ggggcagagg gcagcctggt cgctggattc gggggaggaa ccacatcagg ccatgagctg 960
gagctggtgg tagaatgtac aggagaggcc agccagggcc agctcatgtc agacctcaag 1020
cggggaagat gaatcgagaa tgcacccac gagcaatggg aagccagtct acgatttaag 1080
cagcaaaaat atttccctt ctccaccct gcattcagct ctaccagcac agcctggggt 1140
tctatttca agatagaata gaccagact ccagctctt cttacttc tactactgcc 1200
acctgtcacc cactcatgeg tcccacttg cagcctcgac ccccttcac ctgatctcat 1260
ggcagccagg gaagctccag ggctcgtgag ggctgccatc tcaggaaaga agcaaaagcc 1320
ttcggcacct gcagggcctg ctccaaccac acttcttct tgaccttca gtttcttag 1380
ccactccctt cccacatctc acctgctcc agccacagtg gtgtctctgt gggttctcaa 1440
acacaccagg tgcactctg cctcagggcc ttgtgcttg ctgttcttg ctgggactct 1500
tttttttt tttttttg agacagggtc tctctcttg gccaggctg gagtgtagt 1560
gtgtgatcgt agctattgc aacctcaaac tctgggctc aagcaatct cccacctcag 1620
cctctcaagt agttagctt ttgtttttg ttgtgagtg ggatctcact ctgtgcca 1680
ggctggagtg cagtggggca atctggctc accacaacct ctgcctcca ggctcaagca 1740
attctcctgc ctacgctcc caagtagctg ggattacagg catgtgccac cagcccagc 1800
ttattttgt atttttagt gagacagggt ttacatgt tggctggct ggtctgaac 1860
tctggcctc agatgatcca cctgcctcg cctcccaaag tgctgggatg acaggcatga 1920
gcctgtctct agtagttagg actacagaga ggggcatca tgctgggtga tctcccacc 1980
tttctgctc caactcttc accccactta gcctcgtggc tctctctt acctcttcag 2040
ctctcagtc aggcctgagg acccctgtg aaaattgcaa accacacccc ccaccaccac 2100

cacccactat tgccagcact ttctactcca ttctctgct ttactttct cctttgtact 2160
catcaccacc tgactcatta catgtttacg tatctttctt ctctccacta gcatggaagc 2220
tccaggagag cagagagtgt agttttattc cctgatgtgt ttctgtgcc cgtaccaggg 2280
cctagcacac agtaggtgct cagtaaagt gtgttggatg aacaaataca gtgaaaggat 2340
ctgatctaca ttataaaga aggcactctg gctgctgagt ggggatgaga ctgtcaggag 2400
gaaagaggcc cctgtggggg cctggccagc aggtgggtac aatggtagca gccaggagag 2460
agggcctctt ggactcaagt ggatggggcc tgctcagggc tccggccaca ggaacaaagg 2520
gaagggggcc caggatggcc tgtcatagag gacacattac aactggcca aagttcaagt 2580
caggtttcta aatttgggaa gggatacaga aaaactaaag actctactgg acagtcagtt 2640
attgaaatga ttacatagaa aatgtaccaa gaattaaaa aaaaaaaaaa aagcattatg 2700
aaggggccac cagagactcc cagagaggaa agggactatg ggctggatgc ggtgactcac 2760
acctataatc ccagcacttt gggaggccga ggagggtgga tcacagggtc aggagtcaa 2820
aaccagccta ggcaacatgg taaaaccccc gtttacta aaaatacaaa aaattagctg 2880
ggcatggcag catgtgcctg taatcccagc tactcgggag gctgaggcag gagagttgct 2940
agaaccagg aggagaggt tgcagtgagc cgagattgag ccactatgct ccagcttggg 3000
cgacagagca agactccgtc tctaaaaaa agaaaaaaa ggccagatga ggtggctcat 3060
gcctgtaatc ccagcacttt gggaggccga ggtgggtgga tcacagggtc aggagatcga 3120
gaccatcctg gctaacatgg tgaaactcca tcttactta aaatacaaaa aattagccgg 3180
gcgtggtggc gggcacctgt agtcccagct acttgggagg ctgaggcagg agaatggcgt 3240
gaacctggga ggcggagctt gcagtgagcc gagattgcgc cactgcactc catccagcct 3300
gggcgacaga gttagactcc gtctcaaaaa aaaaaaaaaa aaaaaatta gctgattagt 3360
tgggcttggg ggcggggccc tgtaatcca actactcggg aggtgaggc gggagaatca 3420
cttgaaccgg ggaggcagag gttgcaatga gccgatatca cgccactaca ctccagcctg 3480

ggcgacagag caagactcca tctcaaaaaa gaaaaaaa aagaaagggg ctgtgctgtg 3540
gcctgggacc caaagcacac tactgcaagg tcccaggggtg cctgactcca accggagcct 3600
tgagaacatt catttgcaaa gaatgaatta aaattcagca ctattttatt ctgcaggatt 3660
ccagcacccc aaggacagtc atttttagac ccttcagtaa cgtaataagt aaccggagga 3720
tgtgctgagc ttccacttcc ccagacgggt gcctgtcaca gctcatcagg ccaacaaact 3780
tttcttaggc ctcaaatttg gaaatgttca ctctcagttc gttccttaga tgcaagtcca 3840
tcccaatgaa gtaacagggg ctgagcacct gtccaatctc attgcttccg gggacagggg 3900
cccatgagga tgtcgtttca gcccgggtgac acttgggcaa agtgcctttt ggtttccctc 3960
ccaggctgga acgtgctggc tctgtgaagt tacgtgggc acaagagccc ccccaaccc 4020
ggcaggactg actgctgtgg tcagaggcgc ccctggggct ttgggagcca cagaatcttc 4080
ctgagggcag cgccggagga ggccccagt agagtgccca ctgccaggct cattcctcag 4140
gctgccgag gcctctcccc aaaacaggca atgcttctca gcaacctgcc ccaggagcag 4200
gccaggggag gccgccatcg gcctacagt ctgggctctg gagggttgg ttgtaacag 4260
gccatggttt ctatgagcca gctgggggtg gaaggacaca ggctggattc acctctctgg 4320
gcctcagttt ctgcattcaa aaagtgggaa tcatgatac tgctctattt cttatcttc 4380
agtgtgatg tgaacctcca ataagacttt taaaataact cttctacct tacttttatt 4440
tttcatttat tttaagataa tgtctagctg tctcaccag gctggagtgc agtgggtgta 4500
ttacggctca ctacagcctt aacctcccag gctcaagtga tcctctacc acagcctccc 4560
aagtagctgg aactacaggc atgcaccacc gcacctggat aatttttct ttgagacaa 4620
ggtttcactc tgttgcccag gctggagtgc agtgggtcac tcttggtca ctgcagcctc 4680
aacctccctg ggcttaggtg atcctcacac ttcagttcc caagtagctg ggactacagg 4740
tatgtgccag tacaccagc taatattttt gaaggatggg gtttactat attgccagg 4800
ctggtcttga actccagggt ttaagcaatc taccttccct agcctgccaa agtgcctagga 4860

ttataggtat gagccacccc ccggcctata atcctaccac tttaaaaaag cctgtaattt 4920
tagcacttta aaaaattttt cttaaatttt tatagagatg ggggacagct gtggtctcac 4980
tgtgtgtccc aggctggtct tgaactccta ggatcaagcc atcctcctgg cctggcctcc 5040
caaagtgttg ggattataag cataagcctt accttacctt ttttttga gttgcagttt 5100
tgttcttgtt gctcaggctg gagtgcaatg gcaagatctt ggctcactgc aacctccacc 5160
tcccgggttc aagcaattct cctgcctcag cctcccgagt agctgggatt acaggcatgc 5220
gccaccacac ccagctaatt ttgtattttt agtagagatg gggtttctct atatacctta 5280
attttaagc actgcattca tgtaaattgt gattaacatg gattcaagag agggagtgag 5340
gatgaatgag ccaggcagtc acctcggctg tcacctcca ctctctct cttctgaca 5400
gtcatcgcc atccgttct gcagctgtt gttgactct cctgatcatt ttgcttgcca 5460
cataacttg ctcctgggaa agaatgccct gggcaggccc acatgagtag tgaaaaataa 5520
tctgcagtga aaaataaac taagtagtct ggtccacaga gcagtcttat ttttactg 5580
cagatgaagg agttgacatt caggcttcat tctcattat aagtgttta aagacacata 5640
cagtggattg aacagtggcc ttcaaaaaga tgtatctaca tcctaatccc tgggacctgt 5700
gaatgttaac caagttagga aaagggtctt cccgggtgic attaagttag agatcttgag 5760
atgaggagct catcgtggat tatccagggtg gaccctgcat ccaaggacaa atggctccta 5820
gaaaagaaaa gcagaggctg ggcacagtgg ctcaagcctg taatcccagc actttgagag 5880
gccgaggtgg gtggatcacc taaggatg agttcgagag cagcctggcc aacatgatga 5940
aatcccatct ctactaaaaa tacaaaaatt agcaaggcat ggtggcgggt gcctataatc 6000
ccagctactc aggaagctga ggcaggagaa tggcttgac ctgggaggcg gaggttgag 6060
tgagccaaga tcgcgccact gcactccagc ctgagggaga aaagtgaac tctgtctcat 6120
aaaagaaaag aaaagcagac agagatctga gacagaagag gagagtgaag gaaaaaggc 6180
catgtgaaga tgaggcagag gttggagcca tgcagccaca agccaaggaa tacctggagc 6240

cccagaagtt gcaagaggta ggaagaagcc tcccctagag cctccagacg gagcacagcc 6300
ctgccaacac ctccacctca gacttctggc ctccagcact gtgagataat caactgctgt 6360
tgtttaagc caccagattt gtgtaattt gttatggcag ccacaggaaa ctaatacagt 6420
acctaatttt cacaaccca tcttacagaa aaggaaactg aagtcagaga ggtagtggt 6480
tgtgcagtgt gtaggcat tctgtatta ctataagaa atacctgagg ccgggcatgg 6540
tggtcacgc ctgtaatccc agcacttgg gaggccaagg tgagtggatc acttgaggtc 6600
aggagtcaa gaccagcctg gacaacatgg tgaaaccca ttctactga aaatatgaaa 6660
attagccagg catggtggcg tgcattgta gtccagcta ctccaggagc tgaggcagga 6720
gaatcacttg cggcgaggag gaggagggtg tagtgagcca agattgtgcc actgcactcc 6780
agcctgggag acaagagaga aaccctgtct caaaataat aaaaaacaaa taaacacctg 6840
agactgggta gttataaag aaaggggtta actggctccc ggtctgcag gctgtacaag 6900
catggtgccg gcatctgctt ggttgctggg aaggctcag ggagttttac tcatctgga 6960
aggcagagcc agagcaggtg catcacacag caaagcagg agcgagagag agagagagca 7020
gggagggtgtg cacactttta aatgagcaga tctcacgaga actcaccatt gcaaggacag 7080
caccaagcca cgaggggtct gcccctatga ccaaacctc ccactaggcc ccaccccaa 7140
cattgggaat tacagttcaa catgagggtt ggggggacaa atatcaaac tatatcattc 7200
caccctggc ccccagatc tcatgttct ctcacattgc aaaatatagt catgccttcc 7260
cagtagcccc ccaaagtctt aactcatccc agcattaact caaaaatccc attccaagt 7320
ccaacgtctc atctgaagat gatttcttt cacctacaag actgtaaaa tgaaaacagt 7380
tatttactgc tgagatacaa tgggggcata ggcattaggt aaacattcct gtccaaaag 7440
ggagaaatcg gtcaaaagaa aggggctata ggcccaagc aagtccaaa cccagcagag 7500
caatcattca atctaaagc tccaaaataa cctcttaaa ctccatgtcc catagccagg 7560
gcacactggt gcaaggggca ggctcccaag gcctgggca gctctattcc tgcggcttg 7620

cagaattcag tcccatggc tgccttaca gattggagat gagggcctgc ggctttcca 7680
ggcgcagggt gcaagctgct ggtgatctac cattctgggg tgtggatggt ggcgccccg 7740
tcccgagct ccactaggca ttgtccagt ggggactcta tgtgggcct ccaacccac 7800
attccccctc caatgggaag gctctgccc tgcagcagcc ttctcttg gctcccagc 7860
ttctcatac atctctgac atctaggtg atggtgtcaa gcttcctca ctctgcact 7920
ctgcacacct acaggcttaa caccatgg aagctgcaa ggtgtatgc tggaaccctc 7980
tgaagcagca gcctgagctg tgactatggc ctttgagcc aaggctggag ctggaacagt 8040
ctagatgcag gcaggagca gtgtcctgag gctgtcaga gcagcaggc cctgtgcctg 8100
gacaatgaaa ccattcttc ctctcatcc tctgggcctg tgatggagg gtttggaag 8160
atctctgaaa tgccttgag gccttttg ctctgaggcc tatttctat tgtctcagt 8220
attggcagtc ggctccttt tagttatgca aatctctag caagaggta ctccactgc 8280
ggctgaact cctctcctga aaaagcttt tcttcttg tcataggcc aggctgcaa 8340
tttccaaac tttatgctc tgtttacct ttaaataaa ctctaact taattcatt 8400
atttgcctc gcattgagc ataggaatt caaagaagct gggccacatc tgaatgct 8460
tgctgctca aaatttatg ccacgcttg tggtcacac ctgtaatcc agcacttg 8520
gaggcctagg tgggcagatc acgagatcag gagatcgaga ccatcctgt caacatggt 8580
aaacccatct ctactaaaaa taaaaaaaaa ttgcttgg gtggtggcg agacctgtag 8640
tcccagctac tggagaggct gaggcaggag aattactga acctgggagg cagagggtgc 8700
agtgagcca gatcatgcca ctgcactcca gcctggtgac agaataagat ttgatctga 8760
aaggaaggaa ggaaggagga aggaagaaa tgtctcccc ccagatgtc tgggtcatcc 8820
ctcttatgti caaactcaa cagatcccta gggcatgaaa ataatacagc caaattatt 8880
gctaaggcat aacgaaagt accttgctc cagttccaa taagttctc atttccatct 8940
gagactc acctggcct tggtgttc atatcactgt cagcatttg gtcacaatca 9000

ttaaccagc taatcgggag gctgaggcaa gaggatcact tgaaccagg aggttgaggc 9060
tgcatgagc tgtgatcaca tcatgcagt ccagcttggg caacagagca agatcctgtc 9120
tcaataaata aataaataaa tacataaata acttaagttt atttaaagct gcatctttgc 9180
caccatggag aaaggccagg ccagctcctt ctctctttct gcacgtgttc ctcccacctc 9240
agctgcctct gctcctcaag gaggaacaga gggagtagga aaggccatcc caggaggccc 9300
agcaccccat gacctggctc tggggccttg tgggttatg gattcccagt gctgagtcac 9360
ccctcacagg ctctgtggg caccttgac attgtcaga agcatgtgtt ccccggaac 9420
acacctttc ctgatcatc gggaaggga gcttgtcca gcgaggccac ctgttcagcg 9480
ccacggcccg ccagacagct gcagccacag ccttgccctt gatcagagca aacaccagac 9540
atgtgtgtca tgccccaac ccctctccag gggacacatg tcctttcttg ccaggcctga 9600
gatgaacaag agagggacaa gtccccaage ctctctctcc ttctgctc accactccg 9660
ctgttagatt ctcaagggtg atggtgggct aactaggga accgaccatc ctggttacc 9720
tagaactgag ggggcattt caggaataaa actgcaaaag tctggagcaa acaggagcaa 9780
gttggtcact ctggggctgg tggagtcagg ttctctctg caggccccct ccccgcaagc 9840
atgggtgaa cccaggacag gaacacagag caggccccag gaccgggctt gtcacttaca 9900
agtcttttt tttttttt tttagatg gagtcttct ctgtcatcag ggctggagta 9960
cagtgggtgc atcttagctc actgcaacct ctgccttctg ggttcaagt atccccctgc 10020
ctcagcctcc tgagtagctg ggactacagg tggcaccacc acgcccagct aatttttgt 10080
atttctagta gagatgagat ggccaggctg gtcttgaact cctgacctca agtgatctgc 10140
ccgcttggc ctccaaagt gctgggatta cagggtgag ccactgtgcc tggccccact 10200
cacaagtct aaaccatgcc tcagcacatc aatgccattt acaaaaaggt agagggattt 10260
tccaggcaaa aatagatgaa agacatagga tgattgatca tgtcctgctt aaacataggt 10320
ctgatgctat taagaattga gggctgggag cgttggctca cgcctgtaat cccagcactt 10380

tgggaggccg aggcggggcg atcacgaggt caggagatcg agaccatcct ggctaacacg 10440
gtgaaacccc atcttacta aaaatacaaa aaatggccgc gcgcggtgac tcacgcctgt 10500
aatcccagca ctttgggagg ccaaggcggg cggatcacga ggtcaggaga tcgagacat 10560
cctggctaac acagtgaagc cccgtctcta ctaaaaaata caaaaaaat tagccaggca 10620
tggtggcggg cgctgtagt ccagcaact tgggaggctg aggcaggaga agaattggtg 10680
gaacctggga ggtggagctt ccagtgaagc gagatcacac cactgcactc cagcctgggc 10740
gacagagtga aactccatct caaaaaaaaa ataaataaat aaataagaat tgtagtatt 10800
ttgcagggtg gacaaatgat tctgttctg tggcagaatg ttctcaggag atctctttg 10860
aacttcatg gaaagcatca tgctgttggc aacatcacat ttattttat ttatttatta 10920
tttttagag acagggtctt gctctgttgc ccaggctgga gtgcagtggc acaatcacag 10980
ctcactgcag cctcaacctc ctgggtcaa gcaatcctcc tgccicagcc tcccaaagta 11040
gtggggacca caggcgtgag ccactgcact cagcccaatg taccttcaat atttacattt 11100
ctggcaaagg tagcaaaacc ttaacaaatt ttgaatctag ataataaat tatgaggctg 11160
ggtgcagtgg ccctgacagg gatggctcac atctgtaatc tcaacatttt gggaggccaa 11220
ggtaggcgga tcacctgagg ccaggagtgt gagaccagcc tggccaacat ggtgtaacct 11280
tgtcttaac aaaaatacaa aaaaattagc cagacgtggt ggtgcacgtc tgtcatcca 11340
gtactaggg aggtgaggc aggagaattg cttgaacctg agaggcagag gtgtgatga 11400
gccgagatcg cgtcattgca ctccagcctg ggcaaaagca agagcgaac tctctctcca 11460
aaaaataaaa aaaaaataaa ttaatgaatt aattaaaata aaataaaata atggatagtc 11520
actgtaaaga aaaaataaat gtatatatca gccacaagt gatggaatag agcaccat 11580
ctccctggct ggacagatac atcccacaac acctggaagg cggctccatg tagaactttc 11640
tggactgctt gaggtgctgt gctggagcac ggtgacagag gagctggacc atggacctcc 11700
cccgccccc accaaggcgc aggtccccct gtggtgggtc tgaggaggc atccgtatgg 11760

cctctgcggc ttgggcaggg aatttggggt ccaagtactt ggtgcaaagc ctggaaagag 11820
ggtttgggtg ctgagggcat atcccctggg ccacatgggg gcagaagtgg ggccccctga 11880
agcttggagt cctgggcagg ggcattctatt ttgctgtctg aggccttcag tacttgaagc 11940
aaaatggagg cagaatgtcc cacctaatg cccctgattc ctccaaacca attccagaga 12000
cagcaagggc cagaacaggg atggccctgc ccagggtcat gcancgagga agtggccagg 12060
ctgggatctg aaccaggt aatcccctcc ctgtctctcc tccaggccct caccctgca 12120
tagagccctc cagctcactc atctcggcc agtccatct cctcagctt taaaccccc 12180
cgggatttc ctttctaaa aaacaaaggc ttggccaggc acggtggctc acgcctgtac 12240
tttgggggtg gctcccagca ctttgggagg ccaaggtggg cggatcatga ggtcaagaga 12300
ttgagaccat tctggccagc atggtgaaac cctgtattta ctaaaaaaaa aaaaattaac 12360
tgggcatggt ggctagctac ttaggaggct gaggcaggag aatcgcttga acctgggaga 12420
aagaggttgc agtgagccaa gatcgcgcca ctccacttta acctggcaac agaacaagat 12480
tccgtttcna aaaacaaaca aacaaacaaa taaacaaaaa aaggcggagc gcgatggctc 12540
gcgcctgcaa tcccagcact ttgggaggct gaggcgggcg gatcacttga ggtaggagt 12600
ttgagaccag ctggccaac atggtgaaac cccatttcca ctaaaagtac aaaaatcagc 12660
cagggtgtgt ggtgggtgcc tgtaatccca gctactcagg aggctgaggc aggagaatcg 12720
cttgaacca tgacctggag gctacagtga gctgagattg cgccactgta ctccagctt 12780
ggcaacaaga ttgtttctc taaaaaaaaa aaaaaaaga ctggcccttc cccttcagct 12840
cttctcagg gtcctgagc actctacacc cccgtctaca ctgagcactc caccctgctg 12900
tctacactga gactccacc ctgcatcta cactgaggac tccacccac tgtctacact 12960
ggctgcctcc cgccctcacc tctgctaag gccattcccc gctgcatctg tcttctagat 13020
tctgcagcct tcagcacgtt gggcccctcc ttgtcccct tgagccacct ccagcctccc 13080
cctgagctgc tactctctc ccagcagcct ccaccaagc ccctccagtc cccaagctgt 13140

cccttgcatc cagcactgcc ctccacgtg ccccttcct ccagcttcac agcagggtgg 13200
ggcctccagg ccctgcccac tgtgccatc cacaagttgt ggtgggagct ccgaggggag 13260
gcaggggtgt gcatggactt gggacgtcca agtctgggac caggggcagc tggttggtgg 13320
agtgtggagg gggataggga cttcaggta gagaggctgt aggggcaaga tcgggacggc 13380
ggatgtccct aaggagggt ctgacctggg aaatattgt cagcttctc ttgccattc 13440
ctggagctca gacactggcc ggctctacc ccgcccttc tgcaggacac agtccatcc 13500
cagtgagttc ctagttaga catctccagc agcacggatg ggaaaggaag tcatcaaagg 13560
tgcccaggac cggaggcttt ttctggaggt ggacaggag ggtgtgggtc tcagggtctt 13620
ggctgagggc aagcgtggga ggtcttaggt ctgcaccagc cccgtgaagg cccctcctgc 13680
tccctggtgg agtctagag ggaacagcag cccctaggct ctgacaggag tgggtagggg 13740
cttttctggc ttctactgt gccagcagga tagctgggcc tggcactgag cccaaagatc 13800
acatgccggg gcattggcgc agtgaggaac agaccctgc caaagctggc aaagaagacc 13860
ccatggggtg cagctggtga agctgagagc tcaatgttg ggggagcctg gcaaaagggg 13920
tcctccctc cctctgcagg ccaggatgc aggtttccc tacatgttg taattctcaa 13980
acaatcccat ggccactgga gcaaatgca cagtgggcgg cggcctcggg agcagtggac 14040
agggcacgca gtgccttga tgccagagcc ctgccccaa agtcaacaaa ctctgcagcg 14100
gactttgcac ccgactttg ttccacat acaaggaaag ggacagatca caggccctct 14160
cgctgccctc gctgagccgg aagctgcagc gtgagctctc tcaagccca ttctaggtt 14220
ccccaggcgc acccctgagc cctactgc ctattaagt ctctaatag ccttcaagg 14280
tcttaatgta tgtccattag acagagggga aaactgagc gagggcaagt gactgaccg 14340
aggttcctc gegagcaggg cgtggagctg agaacctgt tattactgct cccacacaa 14400
ccctctggcc gttcttgaa gaaggctgag ccccgggggg gccagagtga cccaaacacc 14460
atgggccgcc tgcgtaaca cgtgcggcca cgaaggggca gcagttccc gcccgccgg 14520

gctctctccg gcgctcagta tccgtcccag gccagaaga agaaactcgg ggaggagggc 14580
ggagggggct gcgtgggagg gcgtggaaga tggacgtggc caggggagtg gcagctgcac 14640
acagtggatg ctgtaagat gaagggaag aacgtgggct ccgagatcac tggacacggt 14700
tccaccttc tcccgtca ctgcatggcc ctgggcggt tgtgaacc ttgaaacct 14760
gttttctt tttctttt ttttgagac agggctctgc tcttgggcc agactggagt 14820
gccgtggcac gatctggct cactgctgcc tcccaggctc aagtatcct ccagctcag 14880
cctcctgct agctgggacc ccaggtatgt gtcaccacag ccgctaatt ttgtattt 14940
ttgtagaga cgggatttc ccgtattgcc caggctggc taaactcct gattcaccg 15000
gatcttctg cctcagcctc caaagtgt gggattactg gcatgagcca ccgacccag 15060
cagagacctc agtttctaa cctgtgccag caggaataat gatagctgcc tagcttggt 15120
gtgctgggaa ttaagtaaga tgaccgggta gcaaatatga agtattactg gacacagagg 15180
gccccaggct gggtagcag cgggtggcag ggctgctgt tctggcctg agctcgaagg 15240
agggccctca ttaccacctg ggtgagctc cgtccaagcc tggcactgt gcgtgggaat 15300
aacttctgcc acccaagtg gcagatttg tgcaaagta agtctgact ctgtggggtg 15360
gacttcgagg cctctcatc ggacctgct ccggtgact cattgcacc tctcctgtt 15420
cctggttaa cacagcccag ctttctcct gctgagcct cctgggcct gctgtcacc 15480
tcgtgccgt gtgctcgca tgccactcc ctgtacctg aatacttgc cctgccttc 15540
caccagctg agagtcaggg ccctgtgag gcttgccca gccgtcctc cgggtttctg 15600
cctctgtga gacttcct gcattgtgc ttctgagagt cccccagcc tgtgagctt 15660
tcaggactgg gacagcttct caggaccgag gcttctggt ctgctgcaa tttaacaggc 15720
gggcacattt tccctggcc aacatcagag actggacatc tgcagatctg tgctagccac 15780
tgagcaccca ggcaccccag caggtagctc tgtaaccaac ccattctgta aagctgaggc 15840
tcagagaggt gaagcgctg gcctggggcc acagcctgcg tcagctgcag agccaggagc 15900

tgagatatgc acctggggct ctgctcacag ggtcctgcac agactgctgc tggagccacc 15960
tatgtagagt caagagagtt catgttaact ccctctcaca tccctcagcc aggggtggggg 16020
ctgacgatag acactcaggg atggcctacc ccccccaaca acccccgta ggtttgccgg 16080
atctccttgg aagaaaagtt ctgggcagaa ttccaccgtt ggcctggcct acactctcct 16140
tagtggctta ggaccctcag cgggtggataa gttgtgggca gaagagatgc aatcaggatt 16200
ctcaccact cacccttgc cagccccaat aagctcaata agctgggctc ggtctgagga 16260
agtgtccagg aaatgtgcaa atggcctggg acagccctgt gtctcttca gtaaggttgc 16320
tgaaggtgag gctgaaagtt ggagaaacag aagccagtgc ttatggttt aattaagata 16380
atggaatgta tgtatgtatg tatgtatgta tgtatgtatt tatgtattta tcttagaga 16440
tagagtctca ctctgttgc caggctggaa tgcggtgaca caatcatagc tcttgcagc 16500
ctcgacttcc tatgccc aaa tgatcctcct acctcagcct cctgagtagc tgggactaca 16560
gacacagcc aactatgcct agctaatttt tatttctatt tttgtggag actgggttct 16620
cactttgtg cccaggctgg tcttgaacct ctgcttcaa gcaatcctcc tgcctcagcc 16680
tcccaaagtg gagggattac aggtgtgagc caccacacct ggcctggaat ttattgtat 16740
tctgcttata aaattaatac attctattg cagaaaagtt tgaaaataaa agaaaggaca 16800
aagaacaaaa agcgtatata atttcacagc tcagatctca ctgctattaa catttttatt 16860
tactttcagg ctttttctt tctaggtaca tatgcagaga ttattttatt ttatttatt 16920
tattttatat ttattttat attttttatt tcattatttt attttatttt attttattat 16980
ttttagagac agggcctcac tctgtcacc aggctggagt acaatggagt gatcatagct 17040
cactgcagcc tcaaacacct gggctcaagc aatccccca ctcagccttc tgagtagttg 17100
ggactaaagt gtgagtctgg ctaatttttt ttactttttg tattgacaga ggtctcacta 17160
tgttggccag gctgatctca aactcctggg ttaagcgat cctccacct tggactccca 17220
aagtctggg attacaggca tgagccacca tgcctggcct aaaatgccac ttttctcat 17280

ttactaaaat cccatggaca cttgacatg tctgtattct atgctattga tctgactgtt 17340
ggcatctaca tcattatggc catctatcat ctatcataat ccattttaac attaaaattg 17400
tgctgctgct tagatttttc tggcctgtct cctatttgta ttctccaga taaattttag 17460
aatcatttta tcaaattccc ctgacagaaa aagccctatt ggatttttgt tgaaaaatac 17520
tgaattttta cattaactta ggaaagggct gggcacgggtg gctcacgcct gtaatcccta 17580
cacttttcga ggccaaggca ggtggatcac ttgaggttgg gagtttgaga ccagcctggc 17640
caacatgggtg aaactcggtc ttactaaaa atacaaaaat tgccaggcgc attggctcac 17700
ctgtaatccc agcactttgg gaggccgagg tgggtggatc acgaggtcag gagatagaga 17760
ccatcctggc taacacgggtg caaccccgct tctcctaaaa atacaaaaaa ttagccaggc 17820
gtgggtgggtg ggcctgtgg tctcagctac ttaggaggct gaggcaggag aatgggtgtga 17880
accagggagg cggagcttgc agtgagccaa gatcgcgcca ctgcactcca gcctgggcga 17940
cagagtgaga ctccatctca aaaaaaata ataataataa tacaaaaatt agccgggggt 18000
cgtggcgtgc acctataatc ccagttactt gggagggtga ggcaggagaa tcgcttgaat 18060
ccaggaggtg gaggttgcaa tgagcagaga tcgtgccact gtactccagc ctgggtgaca 18120
gagtgcact ctgtgaaaaa aaaaaaaaaa ttctgaagga ttgagactct tagactctta 18180
ggtcttcta tccaagagca caatatagct ttcatgtat tcaagcctt ttcaatgcat 18240
caacagaatt ttacagttt ttcatgata tctgtctatt tcttataaaa tgtattccta 18300
gatattctgc atgttttccg gtgtttgtt aataaatatt ttcatttgt cattattcc 18360
taattggctg ttatttgat atatgacatc tgttgaattt ttgattact ttgaaatgg 18420
ccattctttt gtgtttttt ttaactttct attttgagat aatttgact tacagaagat 18480
ttgcaaaaat agtacagaga gtctctgtt ccccttatg ttaaccagt ttctcctat 18540
gttaacatct tacataacta cagaacaatt gtcaaatcta agaataacc tgggcacaat 18600
gctattaact aaactgcaga agctgttcag atctcaccag ttcttctact gctccccctt 18660

tctctccag tgttcaatcc ggaatcctac attatatta gttgtcattt ctctttgggtg 18720
tcttccaatc tgtgacagtt cctcagtcctt tctttgtctt tcatgacttt catttttta 18780
tacttttgaa aaatactggc cggttgtttt gtagaacgcc ctacagtttg gttgcctga 18840
agtttttgt gattagatcg aggtcatgca ttattggaga ggggtgccacc gcctcgatgt 18900
gcaagctcaa tgcatacat cagaggggtt gtaatgtcag ttataccgc cggagaccct 18960
aacctggagc atttcgtgaa ggtgctgtct gccaggattc tccactagaa agttactatt 19020
tttccctttt taattactga atgtctgagg ggaaatactt tgagactatg caaatatcct 19080
gtttctgctt taacttcggc tactaagtt tagcattcat ctatggatct cgcttatagc 19140
aagtattact gtggagtct aatggtaatt ttctgtttct ctattcctt caacctttat 19200
taatatgctt ctctcact tattcatttt gtttcagttg ttataccaa catggatttg 19260
tggatattgg ttattctt tgggttgcaa ttgaatccta tcattatttt gtagtcagt 19320
tgttccatcc gacctgggtc attaggagcc ctgaaattt ggctcccatg ccttttttt 19380
ttttttgag accgagtctc actctgtcac ccaggtttga gtgcagtggc atgatcttg 19440
cttctgcaa cctccgctc ccaggttcaa gcaattctc tgcctcagcc tctgagtag 19500
ctggtattat aggcgtcca ccacctggc cggctaattt ttgtatttt tagtagagat 19560
ggggttttat tatgttgcc aggctggtct caaactcctg acctcaggtg atctgcccgc 19620
ctcggcctcc caaagtgtg ggactacagg cgtgagccac cacacctggc ctctatgcc 19680
attttaacat gcccgtctt tcttttctt tctacttc tgtgactgta agaagctcca 19740
ggatacatti ttgtgccct agacttagcc tcaatcagtt ctcaaaaag ctctggttct 19800
tttatggga tacttagaaa actagctctg tatggcctgg cgcgggtggct cagcctgta 19860
atcccagtac ttggggaggc cgaggtgggc agatcacaga tcacgaagtc aggagatcaa 19920
gaccatcctg gctaacatgg tgaaactctg tctctactaa acatacaaaa aattagtcca 19980
ggcgcggtgg cggcgccctg tagtcccagc tactcaggag gctgaggcag gagaacggca 20040

tgaacccggg aggcggagct tgcagtgagc cgagatcggc agccactgca ctccagcctg 20100
ggccacagag cgagactccg tctcaaaaaa aaaaaaagga aaaagaaaaa agaaaactag 20160
ctctgtatgc tagttttttt tttaagacag ggtctctctt gccccagctg gagtgtagca 20220
gcacgatcac agctcactgt agcctcaacc ttctgggctc aagcaatcct cctgcctcag 20280
tctcctaagt agctgggtct acaggcatgc accaccgtac gtggcaattt ttaaaaactg 20340
ttttagaga tggagtctcc ctatgttgcc tggctggaa ctctggcct caagtatcc 20400
tcctgcctcg gcctccaaa gtgctgagat tacaggcatg agccactgta cctggcctgg 20460
ccaaggtctg tctttttta aaagaagtg ttgtatagtt gttttttt ttattttt 20520
ttctgagacg gagtctcgt ctgtcgcca ggctggagtg cagtgggctg atctcggtc 20580
actgcaagct ccgcctcca ggttcacgcc attctcctgc ctacgcctcc cgagtagctg 20640
ggcctacagg cgcccgtac cacgcccggc taatttttg catttttagt agagacgggg 20700
ttcaccgtg ttagccagga tggctcgat ctctgacct cgtgatccgc ccgcctcggc 20760
ctcccaaagt gctgggatta caggcgtgag ccaccgcgc cggcctgtg tatagtttt 20820
atctcgagtt tttagcgat ttaatcatat tggttacaaa aaaggatgat ttactacct 20880
cctttccaat gttctacat atttttcat ttatctaac tgcatttta aataaactt 20940
taattttaga atggtttcat atttacagaa aatgtgcaa gatagtacag agagttcctg 21000
tgtactccac acccggttc cttatttta tctaacgtg atacacaatt aataaaccag 21060
taacattatt attcactgaa gtccacactt tctttttt ttttctgag acggagtcta 21120
cttctgtcac ccaggctgga gtgcagtgc gcaatctcg ctcactgcaa cctccacctc 21180
ctgggttcag gcaattctgt ggctcagcat cccaagtagc tgggaatata ggtgcccgcc 21240
accacgccc gctaatttt tgtatttta gtagagatgg ggttcacca tgttagccag 21300
gatggcttg aactcctgac ctctgtatct gcctgcctca gcctccaaa gtgctgggat 21360
tacaggcgtg agccaccgcg cccggcgctc atactttct tagatcct tccttttac 21420

ctaacgtcct tcttctgggt caggatccca tccagaaagc aacattaccc ctgccatca 21480
cgtcttcaca ggctccccct gacgggaaga gtctctcaga ctttcttgt tttgttgac 21540
cttgacagtt ttgaggagga ctggtatctt agtctgtttt gtgctgctat cacagactag 21600
ctgagaccga tacatgatac atgaaaaaaaa atgtattctt acagttgtgg aggctgggaa 21660
gttcaagacg aagttgctgg ttggtttggt ctctggtttc aagatggcgc ctgctgctg 21720
catcctctgg agaagaagaa tgcgggtgcc tctcactgca gaagatggaa gcgctaaaag 21780
gaatgaactc ccttgccaa gccattttat aatgggcatt aatccacaaa gcatgaaacc 21840
ctgagaaaca tcaagcttta aagcactggt tctcaacctt tttggtctca ggagcccttt 21900
atactcttaa aacgttttga ggatcccaaa aaaaggcttc tacaggttcc atcttttaac 21960
atttaccata tcaaaaatta aactgaaaaa attttaaat atttattcat ttaaaataac 22020
aaggataaac ccattacatg ctaacataaa tcatgtattt tatgaaaaat agctatattt 22080
atcaaaacaa aaattagtga gaagagtggc atgtataatt tttttgttt atttttgtt 22140
tttagatgga atctattct gtgccccagg ctggagtgc gtggtgtgat ctggctcac 22200
tgcaagctct gcctcccagg ttacacccat tctcctgcct cagcctcctg agtagctggg 22260
actgcagggt cctgccacca cgcccggcta atttttgta ttttagtag agatggagtt 22320
tcaccgtgt agccaggatg gtcttgatct cctgacctg tgatccacc gcctcagcct 22380
cccaaagtgc tgggattaca ggcttgagcc actgcgtctg gcctaaattt ttgtgaatgt 22440
cttaatgcc tgccttctca tattgtttc tgcattcaag ttattgcaa atgttgtgtt 22500
ggttgaagtt tgtaaagaaa atgtggcctc atacagttgt gtagttggaa aggcaagagt 22560
atttgattc tcttttcaa caactatgga caacctgctg ttacaaaacc agaatgcaa 22620
aagttgtagt aaatacaggt taggtgtagt gtggaatctg aaagcatgtg aatgaacttt 22680
ctgagttttg taacattaaa gtccagttgc gtaagctac tgtgatagca tatagcattg 22740
tcctaatact ggaattagta tcagaagtgg ggtgctactg ttaataaata aaaagaaata 22800

aataaatcat gtgatactgg ctcaagaagtc aggcagtagg ctgtgtggaa cctgacatca 22860
cgccatgtaa tacattggca accatttgat ccagctgtct gtcatgatga cttggaaagt 22920
caaccacata cttacagagc ctgtagacat aggggaaaaat agtataaaac agaatactaa 22980
cagtggacct tggttcttgc cagttgcatt tagccaaata ttaaacaaaa gagatattct 23040
tgggcagcaa ctggaccatc ttcaagtaaa agtgaaaggt aataaacaga gtccagacat 23100
ttgtgccc at gcgggttaag aaaaatccag ttgttctag acaccgtata tgaaaacaac 23160
gtgaaaaca agcctttgag tggtaaaggc cgattaacac tcagcgcggt acaaagacc 23220
aggtgggcta acccgaaatg aatgagaag cctgtggtga tgaggaggca gagaagtaaa 23280
atcaagtttg agcatttcgt ttaggagagt ttgggctctg attacttgca catgcaaacg 23340
aactggaaac aaacagatca gatgtctacc acttcttcga gggaattgca ttgccaaaga 23400
agtcatgaaa gcagactcta tactgattag gcattaaaac aaaaacaatc ttaggcccc 23460
taaacttgca tgggcaggaa gtgggctgtc aaagctgttc atcctctaag gtggacctag 23520
ttcctagtcc ccagtataca cttcagatgt ggccctggag gacactggac atggaggacc 23580
tcccagagga tgaggctagg gcttcatttc tccaatgacc tcagctgcct ctattcccc 23640
ttcttctct ggaagtccta tcatcggtat tattattatt atcatcatt ttatttgag 23700
ataaggtctc gctctgttgc ccaggctgga gtgcagtgac atgatcatgg ctactgcag 23760
ccctcccagg ctcaagtgat cctcctgcct cagcctctg agtagctggg agtacaggca 23820
catgccacca tgcttggtca tttttttt cagtagagat agggctctca ctatgtgcc 23880
agggctgac tcaacctcct gggttcaaga gatcctccta cctcagctcc tgagtagctg 23940
ggattcgggt gcacaccacc atgccaacta attttaatt ttttttga tggacaggat 24000
gtacagtgtt agaaatggat tgcttgaga ggcaggagga tcacttgagc ccaggagttt 24060
gatcacactg tgaacctga tcgcaccct gcactccaat ctgggcaaca gactgagacc 24120
ttgtctcaa aaaaaaaaaa aagagagaga gagagagact caaagatagg caaaaaagt 24180

ggaaagcttt atagtggaca aaaaggaacg ctctaagtct gccctattgg catggtgctg 24240
aaggtgggct aactagagat aggggggtact atgtggttga ctatgggtgc atctttggct 24300
ttccctgggt gatcctaagt tggaagcagg gacaaaaatt agggaagctg ttagttattc 24360
atcacgttct ggcagtagtg gactgggtgt gatagaagtt attgttttgg ccaggtgcgg 24420
tggctcatgc ctgtaatcct agccctttca gagttaacg tgggtggatc aggaaggagg 24480
gaggatttgg gaggtcagga gtagcctgg ctaacctggc gaaatcccat ctctactaaa 24540
aatacaaaaa ttagctgggc gtggtggtgc atgcctataa tcccagctac tcgggacgct 24600
gaggcaggag aatcagtga acctggggag gcggagggtg cagtgagcca agatcgtgcc 24660
caatttcac tcaaaaaaa aaaaaagtt atcgtttagc ttctcgatt gttactggac 24720
gtagtaatct ggcttctgc aagtctaact ttcagcagac tggctacatg ggctgtgtac 24780
ttagataag gcagtaagta aagcaaaaat tgatagagca tcaaggataa atagaaaatc 24840
cgtaatcaag cagaagattt gaacacttca ctttcagtaa ctgataaaac aagtagacaa 24900
aaaaaatcag taaggatgta gaagatttga acaacgtaac taacaaactt gacttgattt 24960
acacgtctag aacctgcag aacacacact tttcaagca tactcagaac atttatataa 25020
agtgaccata tgggtggacca taaagcagtt tcaacaaatc tcacaggagt aaaataacag 25080
accgtgtttt ctgaccgtaa gtacagttaa cctagaaatt gaaaacaaaa agctagaaaa 25140
accccatgta tctggaaatt ttaatataca ctttgaaata acaaatggat cagagattaa 25200
ttcaaatagg aatttagaaa tacctgaac tgaaaaataa tgagaatact ataccccaaa 25260
actgtggggt gcagctgaac agtatataga cgaaaagtat actcatatgt gcatacctta 25320
aggagcgggg aggattgaaa gttaatggga ggcaaaagca ggtggatcac ttgaggttag 25380
gagttcaaga tcagcctggc taacagggtg aaaccccatc tctactaaaa atacaaaaaa 25440
ttatccaggc gtagtgaggc tgaggcaaga gaatcgttgg aaccaggag gcagagggtg 25500
cagttagccg cgattgcgcc actgcacccc agcctgggag acagagcgag actccatctc 25560

aagaaagaaa aaaaaaaaaag aaaaggccag gcgcggtggc tcatgcctgt aatcccagca 25620
ttttgggagg ccgaggtggg cggatcacga ggtcaggaga tcgagactat cctggctagc 25680
acggtgaaac cccgcctcta ctaaaaatac aaaaaaatta gccaggcgtg gtggcgggtg 25740
cctgtagtcc cagctactca ggaggctgag gcaggagaat gtcatgaacc caggaggcag 25800
agcttgcaat gagccgagat cgcgccactg tactccagcc tgggcaacag agagagactc 25860
tgtctcaaaa aaaaaaaaaa gttaatggga taaacatcca tctcaagaag ttgaaaagga 25920
atgacaaata aaccaaaaaa aaaaaaatca aaagaagaaa atcataaggt caagactata 25980
aagagagtgg ctgggtgcag tggctcaggc ctgtaatctc agcattttgg gaagcagagg 26040
tgggcagatc acttgagccc aggagtcaa gaccagcctg agtaacatag agagacctca 26100
tctttgctga aaataaaaat aaaaaattag ccaggcatgg tggtagctgag gtgggaggat 26160
cacttgagcc taggaggttg aggctgcagt aagccatgat tgtgccactg cacttcagcc 26220
tgggtgacag agtgggaccc tgtctctaaa aaactaaaat aaggctgggc gcggtggctc 26280
aaatctgtaa tcccaccact ttgggaggcc aaggctgagg tcagcagttt gagaacagct 26340
tggccaacaa gatgaaacct catcttact aaaaatacaa aaaattagtt gggtgtgggtg 26400
gcatgtgcct gtaatcccag ctacttagga ggnnnnctnt ngattatatt ttctcttcc 26460
tacgtcgta ttggactgaa ttcagaatga tgactctcat tggagctctt cctgtctcct 26520
aactacagtg gcttccgacc ccactctggt ttacttca cccctctgct gtcatacga 26580
gtagatactt ccttcttct ttctacttg ttgctcttcc tcaaccccc ccgttgggtg 26640
cccctctct ttatctttt ctgcgcacac ctgcgttctc ttgccctctt atcatccctt 26700
tctcgaggcg gtcctttct ttatccagct taaatacctt ctctctgtt tatttggggg 26760
ttgggtttt atctctcacc ctccctctaa ttctttctt ctttccgcac ccatcaagcc 26820
tctcgtggtt tctcttctc tactctcggg tccccccct ctcccttct tttttctt 26880
accccccaa gcgctttgcc ttttttct ttgcccttta ttcccccc 26928

<210> 7

<211> 29430

<212> DNA

<213> Homo sapiens

<220>

<221> unsure

<222> (4336),(4345),(4349),(4392),(4447),(4490)

<223> Identity of nucleotide sequences at the above locations are unknown.

<400> 7

aggggaaggg ccggctccgt agctcacacc tataatccca gcactttccg aggagagagg 60
atcatctcag gccaggaggt caagaccagc ctgggcaaca cagcaagacc gcatctctac 120
aaaaacttct tttaaagctt aaaaaaaaaa aaaaaagcaa agaggacagt tcaggagaaa 180
agcctgtaga ggcagcacac taaggaggag acgcagccca ggcaccagga ggggctggcc 240
atgggcactc actctccag caggcgagtg cccagcacca gctgcccac ccagacaccc 300
aggacacggc ctgaatggct ccgtattcac gtgggtggtg ataaacaagc aatacacata 360
gccataaagg acaccttagt aatgttacat cataaacgct gcagatcagg gaaatggtgc 420
agggtgaagt ggggtggggg gctgcatgct acatgagaag tgggtcgggg ggctgcatgc 480
tacctgagac agagcaggcc ttgctgggaa agaaggagcc ggcaggcctg ggcaaaggtc 540
ctgggggtggg agcacactgg agcagagtgt gggggtagca tggcgggtgc tggctctctg 600
ggcgcccttc caccacgtca tgtgcccatg tgcccaaggt ctctcgtttc acagccccct 660

gaagctcagg ggtcacagct acacagcccc cagatacctt ggctgcccc aggtcattcc 720
atccagtgat ggacctgctg acctctagcc tgacctctgg gcagcgtaat ttgagaagga 780
ggagaaggga gggcaacaga cctggggcga tgagggatgc acaggggtggc agacacctga 840
ggctgcacct tggagcctca gtctgggtg tgggtggggg atggacaggc tgagggtga 900
agcagctggg cccggccacc atcacacccc aggaccacc agatcacat gaaaaaccga 960
atgtcaactg gcagcccaga gtgcagaaca aaccttcag aaacacggtg gtgactgccg 1020
catcatgaac ataaaataat tacgccctct cccagggat caccctgca ggagttgtc 1080
ccaagaaaca ccagaaagaa ggaaaacgtc tgagtcacaa tatttgctga ggccttatt 1140
gtaatagcaa aaaaaaaaaa aaaaaagaa caatctccag cggcaggggt aactagacta 1200
ttgtctccgt ggaaaggtag caccaattaa ctagtaacaa aatgactgcg gtaacaacaa 1260
aacgttcgac atgtcaacac caaaaaccac acaccagca taaccgtgaa ccatgattc 1320
tactagaatg aatggcagtt atgagaaagc accagcggag acaaagattg aaaaagtaaa 1380
ggtggcctca ttagggagac aagtctctgg gtaatatatt gtaatactgg taaatatata 1440
gittttaata tatttttaa ttccaaattc catatatgtt cctatgaagc tatttctgca 1500
aatattttt tcaggaccgt acatcacaaa ggcaaaaggg ccaggtcagc tctccagctg 1560
agagtgacca cttcagagca gacggcagac tccagggtta gcaagcctgg ctgagacctg 1620
gccccagaca atactcaac ccctctgacc tcaacatcct gtctgtgaaa tggggataat 1680
tactgcacct ccacatcaca gagtgcgagg cttaaacagg atgcttcata gaaaagcgct 1740
caagaggtaa cagccgggag ggggtagtgg tticcattaa taaatgttg ccttcaccca 1800
gccctgggcc agtccaaca caaagcacac accatccact cagactcagt tgcctggatt 1860
caaagcccgg cctggcctcc agctgtgaga ttccgggcag gattcccat ctcccagagc 1920
ctcagttcc tcattcatga aacaggaagt gatcattcct ttattttta tttttattt 1980
tattttgaga cggagtttca ctctagtgc ccaggctgga gtatgatgc gcaatctcag 2040

ctcactgcaa cctcggcctc ccagtttcaa gcgattctcc cacctcagtc tcctgagtag 2100
ctgggattac aggcacacgc caccagccc agctaatttt gtatttttag tagagacggg 2160
gttttgccat gtiggtcagg ctggtctcga actcctgacc tcaggtgatc cgcccgctt 2220
ggcatcccaa agtgctggga ttacaggtgt gagccaccaa gccagttga caactgctt 2280
taaagacacc tctggctgct gtggaaaaca gcctggtagt gcctcaaaaa gttacacata 2340
gaatgatcct atgaccagta attccactcc tacatatata cccaaaagaa ctgaaccct 2400
ctactcatgt atgtacacat acaggtacac gcatgttaac agcagtggtc acaaagccaa 2460
aacatggaaa cagctcaaat gtccataacc gatgaacgga taaatgaaac gtagtctatt 2520
caccacctga cggaggtgag agggggccata aaaaggaatg atgcataaaa acgaatatta 2580
tgccaggta tgggtgctca cgcctgtaat ccaggactt tgggaggctg aggcgggcgg 2640
atcacgaggt aaggagtctg agaccagcct ggccaacacg gtgaaacccc atctctacta 2700
aaaatacaca aattagctgg gcatggtgga gggcgcctgt aataccagct actccggagg 2760
ctgaggcaag agaatccctt gaacctggga aacagagggt gcagtgagct gagattgcac 2820
cactgcactc cagcctgggc gacagaccaa aactccgtt cggaaaaaaa agaaaaaatt 2880
agccaggtgt ggtggcgggt gggtccctgt aatcccagct ctactggga tactgaggca 2940
ggagaaccac ttgaaccgg gaggtggagg tagcggtag ctgagattgt gccactgcgc 3000
tccagcctgt gtgacagaag gagactctgt ctctaaaaaa caaaaacaaa aaaggcccga 3060
cgcggtgtct tacacctgta atgccaacac ttgggaagc caaggcaggc agatcatctg 3120
aggtcaggag ttgagagca gcctgggcaa cacggtgaaa cccatctct actaaaaata 3180
cagaaattag ccaggtgtgg tggcacatgc ctgtaatccc agctactcg gaggtgagg 3240
caggagaatc gctgaaccc aggaagcgga ggttgacgtg agccgacatt gcaccattat 3300
actccagcct gggtagacaga gtgagattct gtctcaaaaa aaaaaaaaaa aaaaaaaaaa 3360
ctaaacaaaa gcaaaaaaac caatgagtaa tgtgtcaag tgaactcat cccaatggga 3420

atgcagataa ttgtttaaa aggcaccatg cacactgggc aggctggctt cccctgggaa 3480
cgcttcttt tgctggatt ccagttggt ttaatcgggc gtagaacact ttctcaatc 3540
cgggattcag gcacctgctc tcagcacaaa ctacgtacac cccgcactct gctgtgggtt 3600
cttggcacta ttaggagaat gtgagggggt gattcagatc tatcttagt ggggtgcatgt 3660
ctgccactcc caggaacgcc cacttctggc aagtcagtgt cagagaaagg ccagctcgtg 3720
gccccctctg cttgagtcc caggaccgt gatcagtcct acccgagca gaatcaggag 3780
tttgaacc caagtccaa caatctcatt ttaacccatg taagcatatc caatattat 3840
atatagaatt cataacagat gtctgggctt ccattccaat agcctatatt ttacactgtt 3900
tatttacatg gttacacaa acaagactca attcaaggta acccaatcct ttgctactat 3960
acaaaaataa gcaacatttt cagtccatgc cttatatata ttcaccaagc attacactag 4020
gcctccaact gtcctcggga gcaagctgca gcctggacac aagctagaga ttaatcagtc 4080
aggaatgatc ctgcgtccag tgccagcatg atggaagaga cagagaaaca gaagacatca 4140
gggctccaga gtcaaggagc ctgcagggtt gttgggcagg atatacacac atacacacac 4200
acacgcacac acaaaaccac ccaagaagaa aagggtggat gaatgcatgg acaggtaatg 4260
cctggagcct ggggatggat aagctgactg cagggtggccc aggcaggctt cctggaggaa 4320
gaagacctgg ctgtangtgg ggtangcang ctttctaaat ggggaaaatc tggctgtggg 4380
tggagtggc angttccga aaagaagaaa agctgactat gggtagacct ggctgttggt 4440
ggaacangca ggcttcttgg aagaagaaaa tctggctgtg ggtggatcan gcaagcttct 4500
tggaagaagt aaacctgact atgggtggac caggcaggct tcctagagga agaagaccgg 4560
ctgtgggtga accaggcagg cttcctagac agaggaagat ctggctgcgg ttagagtggg 4620
caggcttcta agaagaggaa gggctgactg tgggtagacc tggctgtggg tagactgggc 4680
aggcttcctg gaggaggaag agctggagca ttgaaaaaca aacatgactt ggtgaatgtt 4740
gagcatgcc aggcctgatc ccagaggca attacgcact caagtactt aattctactc 4800

acaatgcctc acaaacaact tctctgacac ctaacacagc tctgggcacc ttctagcttc 4860
agctcctcaa agcagttatt cagctacta ccctgcacac ctctcacac cccaaccca 4920
gggacaggag ttctgccaga tgccaaagct cctgatgcca aagcctgggt ctgctccgg 4980
gctcctcttg gtctaactgt ccaccccgca tcggcatgat gtgcaaaaac aaggctttgc 5040
aatctgcctt gatgcctggc ggagcgagtc cctcccgatt cgtctcttc agaaacacct 5100
gggctgccct ggtcctgtta taccccaac acattctaca gtcagctccg caagttccac 5160
aaagatcaac gctggcggtt ttatggcatt ttattacag ttttacaat ataaaaaagg 5220
aaggatgcca cagctcagcc agcaggacag acagagatct atgatgcttc tgctgcacca 5280
ttgtttgtgg tcaagaaagt ctgtttcaa tgattatta aattgtggtg ggagatggat 5340
ggtggcagtg gttaccagca acatgaatgt tcttaatgcc actgaacttc acactacaa 5400
atggttacga cgataagtgt tataatgtatt ttaccacaat taaaaacagg taaatgcagg 5460
ccgggcacgg tggtcacga ctgtaatctc agcactttgg gaggccaagg caggcagatc 5520
acctgaggtc aggggttcga gaccagtctc gccaacacgg tgaaactctg tctctattaa 5580
aaatacaaaa attagccaga tttggtggtg catgcctgta atcccagctt ctaggaggc 5640
tgaggcagga aaatagcttg aaaccgggag gcagagggtg ccatgagctg agattgtacc 5700
attgcactcc agcctgggtg acaaaagcaa aactctgtct caaaaaata aaataaata 5760
aaaataggtg aatgcaaaca tatggtatag taataattat ggctattatg agctacaaaa 5820
aagaatgact tgggactaca gttacagccc tcattcagga atttgttta aatgtgggtt 5880
ggtcgctaag gcatgtacac aacatttga cgttcaaata ttctagatt tggacagtga 5940
gcacccctct aagctggctc ttctgtcca gaggtccca ccagtcctcc agaactctt 6000
tgctttctta cacaataaga tgcccatgc tcggcttgta ctttcttg cccagccct 6060
agaaccagct tcttcgtgga caagctctga ctctttggg tggagaatgg tattcagaaa 6120
cccagacctg ggctctggtg tgctcactgc tacttggggt cattgctctt aggcctctct 6180

gctgatggag gtaggatata cacgtacagt ctccctctt cccagattcc gtacttgagc 6240
tcgcctactt gctaacattt atttatatcc cccaaattaa acctcacagc acttctgcaa 6300
tcactcactg acttgacagag tgtgaaaaaa ctgagtcacc atcacacgtt ccaaactgag 6360
gtcaactgag gccacaacgc cccatcttct tgctccggct gtcgagatgt aagcaagtgt 6420
ccttctctcg gtctagctag tgccatgctt tccacatcac tgtgctttt gtgggcaatt 6480
ttgctgtata aaatgtcccc tgcacatatg ctgctgtgta gtgctcctag gtgcatgagg 6540
ctgccccacg ccttacagag agaatatgca tgagaggctt tattcaggta tgagttatag 6600
cgtagttggc catgaattca atgttaatga atcaacaata tacagtaaata aaggtgcttt 6660
ttagagacag ggtctcactc tgtcaccag gctttagagt ccagtgggtg gaccttggt 6720
cactgccgcc tcaacctcct gggctcaagt gatcctcca cctcagcctc ccaaactgtt 6780
gggattacag gcgtgagcta ctgcactcag cctaaataag gtgtcttaga aacacacata 6840
agacaagggt atgggctgag tgcgggtggct catgcctgta atccaacac ttggggaggc 6900
caaggtggga gggtcactg aggccagaag tttagacta gcctgggcaa catggcaaga 6960
cctcatctgt atatttttt aaatcagaca ggtgtggtgg tgcctgccta tagtcccagc 7020
tactggagag gctgaggcag gaaaatggcc tgagcccagg aggtcaaggc tgcagtgacc 7080
catgattgta ccactgcatt ccagcctggg gtgacacagc aagacgctgt cttaaaaaaa 7140
aaaaaaaaa aagccaggtc aggtatcgaa cagttggcaa aaacgttgtg acctgaggct 7200
cacaggaacc tagcccgatg ttcccctag gagcaatggt tcagtattca ataattcagg 7260
gttcccagtg actttatgga gcataactt caagaataac aagaaccaac tgtactgtg 7320
tatgtatact cacacttta tttatttta tttatttt tgagacagag tctcactctg 7380
tcaccaggc tggagtaaaa tggcgtgac tcgactcact gcaacctccg cctccaggt 7440
tcaagtgatt ctacgctcc caagtagctg ggattacagg tgtgccccca caaccggcta 7500
atttctgtat tttatgaga gacggagttt cccacattg gccacgtgg tctcaaactc 7560

ctaacctcaa gtgatccacc cacctcagcc tcccaaagtg ctggaattac aggcatgagc 7620
tgccgtgcct agcctacata cactttata cacacatgca tctatgacta ttctctatt 7680
tctgtgatg tgtgctggc agtacctaca gttcagcta tgtgtctgg tactgtctcg 7740
tccaagttg taagcacctt ctccaaagtg caaagcctgg ctgtgttac tatccatag 7800
tttacttatt tgctcaatca atttacttat tagctccata accagcttcc catctgctcc 7860
agtagcctct gctgtcagtc acctctgcac cctacccac ctgtctccg gatgctggat 7920
gccaatcacc cccgacacct ctacatagca ccacctcga catgctgctt ctttatttct 7980
tatttattg ttgagatgg agtcttactc tgtgcccag gctggagtgc agtggcacga 8040
tccaggctca ctgcaacgtc cgctcctgg gtcaagtga ttctctgcc tcagcttctc 8100
aaatagctgg gattacaggt gccaccacc acgcccagct aatttttga ttttagtag 8160
agatgggggt tcacatgtt ggccaggctg gtctcgaact cctgacctca agtgatccac 8220
cttgccctct caaagtgtg ggattacagg tgtgagccac cgcgcctggt ctgctcttt 8280
aaatgccagg caccaacatt tgtcaatgg ggtgggagga aagaacaggg aggagagcac 8340
actgccggcc cctgcactga atccactgat caatctgggg gcaactgcca tctccatctc 8400
ctgtcttct atccgtgaac atctactgca gtctctcca atgtcttct gtaaagtgt 8460
attatgttt gcatacaggc ctgcatatt agttctcaga tataatccat atactttata 8520
taaaattcaa accacattta aaaaaataaa actagcatga ctataacgga gtctgcaaca 8580
ttctcacaga ctttatgata aaacatgaaa ctcaaagat acttaggggt gggcaggagc 8640
aatgtttaag gctgcctgga agctcctcca tccctgagcc agaaagtcct atctcccctt 8700
caaggggaaa tgcttgaaaa agcactgatc aggctaaaat gacagggatc agggagtaat 8760
caaagtacaa gtgagctggt ctctccatt ctgagcacag caaagttcag tctctccaag 8820
tccaagaatc atacacctgt ttgccaagaa tgaagttcag gtgtctacaa gtggctgaaa 8880
atattcattc ctggccatt aacaacattc ttggcaaac cataccttag cttctctgg 8940

aaatttctta aggtagaaga aacaggaaac acccaggctc gcttttatgt agacagtcc 9000
atgaagccag ggaccttccc cacatccacg tttaattac ctgcacgcag ctacagtgt 9060
attcaacatc tacgcgtctc tctactggg gtggcggtgg cactcaaac cctcatgcag 9120
ctacgatgac cgcaatttg gcaacataat tcatgttt tcttgggct ttaccaag 9180
tcagtgcac aattctgcag ttgtctaaag attcaaatg agggactga cattacaac 9240
aataataaaa tcttgggtt ctttaacca agcacatgt cgccttta gagaaagctc 9300
tgcaaaactca agctggagt ggatactgc tgacatctc aagcaccca ggaatagctc 9360
tactcccca ttccacctt ggctgaacca tctatctcc accaattccc ccaacatccc 9420
tccatcgtc catccatcca ccaaggacc tgctaagcca ggaggtctc ccatctacc 9480
ccacagcctg gcctcagccc acaagggtc tctctacatg aatcccaccg caccagagta 9540
gaccaagtct ccgtagact ccacctgac cacctccatg cctccagcca ttccacccc 9600
taaaaaccct ccctggctc tacaccagc tgatgaatac ttggctgaat gtgacctggc 9660
ctctggacc cagggaagc ccacgtctc cgtaagccc ccagctcacc ctgcctctgc 9720
accttactg gagagagccc gcaattcacc tctcagggc aggcattggt gatgccacc 9780
agtggaatct ggtgcaaagc agggcccgt gcagagcagg gctgcctgca gagcaaggcc 9840
ctgggtgctg ggccgagcac ctccaatgt ggccgtggaa ccatcctcc cattccaggt 9900
gctgtctcca tcaagaatga gcgagctgt gacattgca tgacaataat gaataaatac 9960
catatttgc ttaaatcca gaatagatgt ggccagggtt ggcatatgac tgttgggaaa 10020
ggacagttg cctcttccc aaccaactg gattataaaa agctttctt aacgaccaca 10080
agagcggagg agctcagggg cagacaaaag gaaggctggc tgcagaaggc gggagagtgg 10140
ggcctcagg ggcgggtgg gagagagaaa gcctggagct gcaccccaa ggtctgtgta 10200
catcaggtgc tacagaataa caccacctt tccagcttg cccacactg cccttccc 10260
gcccagtcac ccaagacga cccactccc cacacacacc tcacatctc ccgctcaca 10320

ctcaccagct tcggctctca atgcaacctg gaacctgccc ttggcctctc agctcagcca 10380
ccccattcc tgttggtccc ttggccccca tcgaattctc tctaattcta atgcacacac 10440
ttgcacactc aaacacacac acacacacac acacacacag cccagaggaa aaccataatt 10500
gactgaggtc caggcaagtt tcccagagcag ggaccacatt tcaaaggta gggagcagg 10560
cgaacaggaa acatacaggg ggcacgttg ggggtggagc aggaaataag aatcacttg 10620
caaaagataa aaagaaaatg aggtagctgg ttccagacac ctccgagcac acagaacagg 10680
acaggcgcct ccgggtcttc cctcaacagg gagatgggcc aggcaggctc ctgtgtctcc 10740
accgcagagc tgggggctat ggccctgaca ccaaggccct ggggcaggcg gggaggcagc 10800
tgttctctg cctgtgtctc cgggcagggc ctggccccac aagggaactg gccgaaggct 10860
ctgttggtc actccggaaa gtcctgggag acaagcaaag gacttgctag gtcactcaa 10920
acggcccaga tgtgacaact gtgaagaagc cacaccaaag caaggtgaca gaacaatgtt 10980
ggtgacgtca gggtatcagc ttacgtcaa ctccacttac ccggactcac ccgtaacctg 11040
ccgtctcttc ccaaccagta aaggatgcct aggtagaggc gcacaaggcc tggagcataa 11100
ttaccatgtt aaaggctctg agaagtcctg cggtagaggaa gcctagtcca ctttctctc 11160
cctaggattt cccaactgcg cctgatcaca gaacattttt tcatttcac tcaggaaaca 11220
tatttgaaa aacactggcc tagaggcaga agtgaaatgg aaaacacaaa agtaaaactg 11280
aacaggaggc actgggcaga gaacggtag aggcgccctg aatcctggac cggtaggat 11340
ccccagcttg gcatgctccc ctccctgggc ccagaccgcc tccccatt tcttgataa 11400
gaaggcta at gcgcatcagg gtgaagggtc tgctgggct acacccccag gctcggcca 11460
accaatgcg gctcctgca gagccagtga ctttctgat ttggctactg tggaattgtt 11520
tgcaactaac cccccagat acagatacaa atgacaggat gatcagatgt aaaggacca 11580
caggtctctg tgatacggct tcatgcagcc agcatggcta gtgccgtgca gaatgagaat 11640
gaccccaggc aagtcctgc ctcccagacc cagaaccca tggagccac cagggtgtgt 11700

tcacaagcac tgtctgggtc gggcagagat tccagcaaga ggagggaaca tccatgcacc 11760
ggagccagtt accagaagca aatgcctct tccaaaacc aggctattaa tggagtccac 11820
tgttgagtgg agctgggtc tagctatgga atactgcaca gcagagatct tctgagaga 11880
aagcagtttt ccctgaaagc catgtgtcct ccactaactg tgtttaatt gggcgaacgt 11940
ctgtatctca ttgcagtggc cgcgcatgtg ctgacaagg gctgggggcg ggggtggggag 12000
cagaagctca ggggcctggg aggggaaggaa acaggccacc agggctcccc agaaggcatg 12060
tatctctctc acaaacacac gcatgcacac acacgtgcac acatactctg caagccctga 12120
gttagcaact gtggaatgtg accagctcag tgatcccagg acaagctgct agggaatatg 12180
acatttgatt gatgtctgca aatgtgcgtt ttactaatt agaaggttta gggcagagca 12240
gagaaaaata tgtatttcag agtccagtt tgacctgcca gaaaccagcc cattactaac 12300
attcttattt tcaacaaaat atagcattct gattacatac catcttggtt ccacgcctcc 12360
tgccttgcca agccccgga agcggcccaa ggccatggca aatagtgaga gaaacagttc 12420
cagggtggag actgactcag ggggtgcagt cagtggggcg ctgatggccg gtgggaggcc 12480
agcagtcac accctctct tgggacagtt gagtagctct ccccagggt catgtggcca 12540
ctcaggttca tatgggaggc gagaggagtg gcagagtcca ggagagtggc tccgaagtca 12600
ctgtccctc caggcctcag tgtcttcac cattaaatgg gtaggctgag gtctgggatg 12660
acaaggaggg ctgcaacta ctgaaacca tgggaggctg ttcgccgatt tctttattg 12720
atggaagaaa acactcgtat aattcaagta ccaattaaaa ggcaggcact ggaaccaccg 12780
ctgccaat cctagtttg cctataccaa atttgagcaa gtaattgac ctctcccagc 12840
ctcagtttct tcgtctgtaa aatgagggtg gggatggccc ccagcccaca gggcagctgg 12900
aaggattaaa gaaatcaaac atctcttaga gccacctgg cacactgtga tacacaacaa 12960
atgtagcta ttttgtcta tgaagtctag attttatac ttgggtgttc taaagcagga 13020
tacatttatt taaaaacaag gattttcatt aaacacgtac cccacagaca gcaaccccat 13080

ggagactgct cttaattcag gccagtatcg aaacgactct aactacaagc ttatacagg 13140
tctcttggt gtccttcaaa tccaactaag gtggtacttc tgaagcactg tgcacatgtg 13200
tgtgtgcatg cacacgtgtg ggaagggcgg gctcacggat ccctcaggta cccacccac 13260
gcagtctcaa gtcacaaagc gacagagcag ccgaggaagg tctgtgcccc actggaccct 13320
cgtgaagcca ccaactctac ctctgcgcg gtctctgcag actgggctac ccttggtg 13380
gggaccagca ttgatgcaa gaaaggcaga cagaaaagga aaagggaag ttcgactcca 13440
gataacacag acagtaccaa gcccagggt ccataaatgc cacgcagatg gaagcattta 13500
ctgcgaggcc acacagcaaa cgcacggatc cagggaagg ggtgcagact gcggtgcccc 13560
tgagccatga ccctgcaaat taccacatg ggaaaggagg ctgccaacc cccgacagt 13620
cggctgggt ggcacagact cgtggtttcc atcgaggtgg gaggaggtgg gacgtcccag 13680
cccccccc atgcccactg cagagggaag cgccgtttc ccctgtgtgg ttacaaagg 13740
ctcattgtt ttctcacag ggaggaaact ggaggaccga gtcagaacg catttagaa 13800
ctggcagaaa agaacatctg gggaaggaaa cacatticag aaacaaacat acctttgtac 13860
cagcttttat ttctttaag tgttgaaaaa ataataataa taaagacatg ccaaatttat 13920
catcgctcta caaatccct ttattgagca aaacgtggca gctctactt caaatgatta 13980
ctgttcctgg aaaattgcag caacgtggat gccaaagccc gaaggccgcc atcagcagcc 14040
aaacaaaaga tgccacctcg ggctccgca cactgtacca tgccaggga ctggacagat 14100
ttggggaatg ccacggttg ccttaaccc ctgcccct ggtctcctga tgcatctcag 14160
aggctaacat tcttgagga actggcatt ctagttgta aatatgcatg tgggttggg 14220
agctgcctgc aaagtccagt gttgacgatc agcttgatt tccttggat caagttacg 14280
tgtcgagtct ggaagttaag aagaatttg agaagctgag cactatggtg tgcaggccc 14340
tgggtgaact ctccaccaa gcattcattg tggactgaca gcgtgcgagg ggctctgcag 14400
gcaggtcac aggacgaaac acattccctc cgggggaaac ctgcaggaaa gctccctctt 14460

cttcctaagg tgccgggcct agcttcatgg gtccctaccc tccacgcctg tcacactttc 14520
tgagtctcat gtgggagctg ctctgggtc ctgacttcac tcagtcctca taggaggtgg 14580
aactactgtc accccatttt acagatgggg agactgggca caaggggacc aagaaaccaa 14640
tgcaaagtca cacttggtgg atcagtgcaca ggggagatca attcccaggt tctttctgca 14700
agagttaaat tgttttcatg ctgcctaagg gggggcaact gaaagaccac tgcatatctt 14760
tgccaaaagg gtcaagcaca ggagccgcag ccagtgggtc agatccgcag aggcgctggg 14820
gtgaccctcc ccatacctgg agggatgctt gtcccctcct ggccttcaact ggggtcccctc 14880
atgaccgtgg cctcccagga cctcagcaca atcccgggtcc tgtgtccag gacaagccct 14940
ccgtcccaa gactgtgagg aaatggaacg aagaggggct cgctgcagcc cagcacccac 15000
actgcccctt ctgaggggca agaaccgtcc tggaggactt ggctttggag ggggagcctg 15060
ggaggccagt aagtcaaca gccttactg ctcatgggtg ggtatccacc gcaggcccc 15120
acctgctggg gcgggcaggg acgggcggca cagcttgcc agggcagata acccccacct 15180
tggccagggc gaaggcagga cacgtgggt ccagcctggc cccaccatcc ctgcacaaca 15240
ctgggcaaag tccacgttt cctcaactgg gtgtgacat ctgcaggaca ggggcatgga 15300
ggtacagagc gctgaagcca cacagcaacc taggagcgag actccatgcc tcccggggga 15360
ccctcccca ccatgaggac catgaaggct tccatgtgc cgcaaggact ctggtgtgga 15420
gacacacgtc tctacacag ccaggcctaa cgctcttga actgggtggt cccacctggg 15480
ctcacagctg gagggccagg agctcaaggc ttcgcagggt ctgctctcat cccagaggcg 15540
atggggagcc acagcaggct gcaggagaga ggggtggccc cctccactc agaggcccca 15600
tctgcccac agactggaga gcacatctt cagcaaccac ggagcgccaa ctgcgcacag 15660
ggcctggtcg tcagagcggg gcaaaggcac tgaccgtcac ggccaggcg aggggaagacg 15720
ggtgggcagg gacctgggc agagggggaa gaacctggtg cccaggctgg ccctgccttc 15780
agcagtgaag ctgagtgggg aggcgctgat gcagggggcc agaaagggt gctggtcagc 15840

cgaggaggagc cccccacaga ggaagcagcc agcccagacg cagatggcag ggtccctca 15900
acaatgtcct ctgaaaagga gaggcgggga ctgctctggt gacacctaca aatagatagt 15960
cagccctcag ccccttgcca tacttctgac aaagcagagg ccccagggg aggcgcaccc 16020
gaaggtacct gcacctgtcc ccagactcc tagagccac ctgaccccat ccaccaggg 16080
ctccagctac aaaataaatg ccgaggccag ctaggcaagg acgcacactc ggtaccgact 16140
gaataggctc cacgttgtca tgagcgcaac ccacaggcca ccaggccaca ctatgcagag 16200
ctgagatggt ttcggccaag cagcctctca gctgagctga acaagtccag agtccccggg 16260
gggtcgtcac tatggagtaa caattgcgat gcgatggtaa ccctaacagc taaccgtcac 16320
tgagccaggc cctgagctag gtactttca acgtgcctc tctgcagcct caggacgagc 16380
ctgtgggagc ataaagatca ttccatatca cggatgggga aactgagctc tgaagcagtt 16440
aacgtgcttg tcccagaccg cagagctagg agcaggacac aacagcaggt caggcaggaa 16500
cggttgaggg gggcctgcat gggcttctct ggaggctgcg catacacga acccccagga 16560
ccccgacct gcacctgcag ctgctactg cccctcagt gactccagca aacctcgggg 16620
taggggaagg aggttgggaa tacctcgggt gtccgaaaca gcagcttctg cttggaggcc 16680
actgctgcat aatggttct gcccagcaca cccaagcca cctgtgccac ctgtggtgac 16740
ctccagcat gccttggga ccaagctggc cttaggtgct gtgggcagcc aagaatagaa 16800
caggggccac cctcctctt cactaaca caaagcaaga ggcgggcact tcgactgagt 16860
gcacccctct agtcaaggg cctcacggt cacaggggtc agggcaagat cccaattctg 16920
cattcccgtc tgccttcat cctgctctgc caacaacagc cagtgaggct ggggacatcc 16980
ctgaacctgt ttctcacctg aaacacatca taccattgga cccagccct ccgggagagg 17040
ccctaattccc tgactgtggt gagatcagat cactggttaa gtaccagaa gggccttggg 17100
caggggctcc aggggtgggg ggtgatgggc gtggtggtat cccgctctgg gctatagtc 17160
acctgatgg aggaggtctg tggtcagaac cgggctgtgc agggcacagg agcccagagg 17220

gacccccaga gctcacctgg tggctctga gcagggtcc ctcaaccctc agagaaaagc 17280
acagcaagga ggccgcccag agcccagcgc ctagcaccca gtggcgtgcc agacctgcct 17340
ggatcctgga gatctctcat caccctccaa gtcagtcag cccaaccag ggaccacag 17400
cccacggggc cgtgaagggtg tgctgagtc aagaaggcct tcgacactgg gaagccaagt 17460
ggcacctcct ggtgtggagc aggcggaatc ccaccagcct ctgctctgcc agtgggcaca 17520
gttgagcat gagcagaagg ggctgttgc taataaacgt catttcctta agaggataaa 17580
acctttcaaa acagatggaa atttttttt aattaaaact ggtggccaaa gagatggaaa 17640
gcaccccttg tgcctccctc ccacgtgac ccacctctg cacacctcaa gctgttcgct 17700
gccaggtgt ctctgaggc actgggggcg ggtgagaatc cgtgagccct cggccagccg 17760
tggtctctg gagctctgcc ccaggccatc agggcacacg ccgggcaccc tgggggccac 17820
acagggcaga gccagctgg gtcagcacac agggccacac tgggcacaca agtctctgag 17880
cctccctgt ggacgcagct ctactatcc cacccta ggtcccgagg atctgtcca 17940
cagggtgata tgctgtcaca gacctacc agagccatgg cctgctgtc cggcgagc 18000
caggtagtca ctgtccac agggacaggc aacccgcac ttgggggctg ctctcggca 18060
ggactagagc tccagcagct cagccctcct gagaaggaga actccatgct ctaaggga 18120
gacgcagcg acggcaccaa agccaccaca agccacggg gccctgcatg gcaggtcagg 18180
agtcctgac cactcgtct ttgtaaccag agctgcagt gagtctacga ggcaaggact 18240
gtggcgggca gtggccacag caaatgaatg agtgtccaa gggagcaggc ggctcgggg 18300
aggcacagc gggaccagg agtctcgg cactgcagca aactccctgg gcccctgag 18360
cagcgaccag gtggcaagt catgaactcc cgggggcata acctgggagg gtgacactct 18420
cttcgtgtc aaattcttga gaacgcatia aaaatatcac tcagtcacct actctatagt 18480
tttaactcaa aagtacaaa gtagccaggc gcggtggctc acgcctataa tccagtact 18540
ttgggaagct gaggcaagag gatcactia gccaggagt tccaaatgaa cctgggcaac 18600

atggaggggac cccatttcta caaaaaaagt gttttaaaaa attacctggg cctggtggtg 18660
tgtgcctgta gtcccagcta ctccaggaggc tgaggcgagg gaaccacatg aaccagggg 18720
aggtagaggc tgcagtaggc tgtgatggca ccactgcact ccagcctggg taacagagtc 18780
agactctatc tcaaaataaa tttaaaaagc accaagccag gcttggtggc tcacacctgt 18840
aatcccagca ctccaggagg ctgaggcaag tggatcacct gagtcagaag ttcgagacca 18900
gcccagccaa catggtgaaa ctccatctcc actaaaaata caaaaattac ccaggcgtgg 18960
tggcgggtgc ctgtaatccc agctactcag gaagctgagg caggagaact gcttgaacct 19020
aggaggcaga ggttgcatg agccaagact gtgctactgc actcaagcct gggagacaga 19080
acgagactcc atctcaaaaa ataaataaat caatcaaac caccaagact tttaataata 19140
aacatttatt attccataat tcctttttg catgattaaa aatgtttata taaagtttcc 19200
tgaaaatggt aagaatgcc aagtaaggct gcaaatgccc aagccccac cgtggcatct 19260
cacggagtct gggccctagg aggtggtgg gtaccacgtg gaccgagac ttcacagtca 19320
agtccctttg gggtagctg ggtttccac accccagaaa tatgggctct tactgcagga 19380
ccatgggggt cctcacactt ggcccagaag ctgtcacata gccagacagg tgttctacaa 19440
cctaggctag agggagctca tgctccagca gaattcgagc cagaggagggt aaaagatggg 19500
taagatctgc tccctggaca gatgaggcct tggcctcaga acagttactg atcatctacc 19560
agacatcaca ctagaggcag aggggcgcag acgaagacag cccctgtcct caaggccctc 19620
ccaggttggg tggaccatgg aaggttcag acagatctgg caagagaagt gccacacca 19680
ggggcagaag atgggcaggt ctgtcaggg cggcacggcc tgccaggcca aaaagttcca 19740
acttcagatg ctggagaatg ggcacgactg tctgagaaag ggaaggatgt gatgaaaact 19800
acttgagaa aaattaatct ggccagagca taagataaat gggcaaaggg gaggttcag 19860
aaagcaagga gaccaagtaa aagctgatgt cattggctct gaatctaggc ttactgaa 19920
tatgcaccgc agggcctgta ggtaaagcct cagagcccag ggagtctgag tggaggagag 19980

ggcagggggac agagctgggg cctgtgtcta cagtgtcag gaggaatagg catggacgtc 20040
agctcggagg ctccagctga agtgaggagg cggccagggc agcacggcca cgcccggatc 20100
cagactcctt tigggaagca agttcgtctt gggggaaagt ttggagaaat ggcctttacc 20160
cgcagaagca agccccagaa catacttgc tccaaaacta tctcgtacag tgaggacgtt 20220
aagcttcagg tcccctagag gagacagtct gctcttctt ggggcagAAC ccaaggtggc 20280
cagagcctgg aaggcaccca gcaccaggc tgggtgttc cagcccaggc cacacgtca 20340
gatagctatt aatgccccgt tgagcaattt cctgagagct ttgccaggca ggtaccgcct 20400
ccccatctga actaatacag ggtacatcc caaggaagaa atgaaagggtg cccacatttt 20460
gctctgggat taactagga ggggagtgt aattaactca gtaattatat ttgccatcg 20520
gctaagtcta aaattagtgt gcattagaat ttcttctg agcagacacc ggagtgtgtt 20580
gggcagcagg agtggctcgg gcaagtggc acaaaggga cctccagagc ctccacaaa 20640
gtcagcaaa accacaaat gtcaaggccg gctccactgc accagcaga tgaattcact 20700
tccacagcct gagaccgcca gctcatcgga ggccatttaa aatccagccc tctgacacct 20760
gctggatc accatttacc gtcccagat caagagatca aagggtggaa cctgatagga 20820
cggctctgaa gttcaccaca aaagcataaa cgtgcaagca gagccaatac gtctttgaa 20880
aaggacaatg aggtgggaat ttacataact gatcttaaaa tatgttctga tgcttcagag 20940
atggagacag cagcattccg gtacacaaag acactcacag gcagtggagc acagtgaagg 21000
gtctggaatc aggaccagg tgtctgtgga cactacacat aaaagagcag catttacaat 21060
gaatggatag gatggaccat cccaccaagg tgttggaaca ctccctattc actggccaga 21120
cccctacctc ataccatata caaaaaaaaa aaaaaaaaaa aaaccagac agaataatgt 21180
ctgaatgtaa aacataaaac agtaacagtc ctggaagaaa ataattggagg atatattat 21240
aatctggaga tggagtaaca agggatagga aaaaagccat agggaaaaag tagagttatg 21300
attatatgaa gcttcttaat atctttatga taatgtacca ccagaaacaa ggatgaagga 21360

ctagctacag accagcagtg aaacctgaaa caaacagAAC aaagaattaa agtccatacc 21420
aaataaagac ctcccacaaa tctataagaa aaagataaac aggctggcac cgtggcttat 21480
gtctgtaatc ccagcacttt gggaggcgga gatgggtagg tcacttgagg tcaggagttc 21540
gagaccagcc tggccaacat ggtgaaaccc tgtctctacc aaaaatacaa aaattagcca 21600
ggcgtgggtg cgcatgcctg tagtcccagc tacttgggag gctgagccag gagaacagct 21660
ggaacccggg aggcagaggt tgcagtgaac caagatggca atcgcgccac tgcactccag 21720
cctggaggac acagcgagac tctgtctcaa aaaaaaaaaa aaaagaagaa gaagaaaaaa 21780
gaaaagaaaa agacaacaga aaaatgggcc aaggataagt gtaggcaatt tgcagaaaag 21840
taaataccaa taaaccagaa atgagggttg tgcaaatcaa aagggttat aattttaac 21900
caaactggac caaagaaaac accaaaaacc aaaatcttgt aattgccagc atcagagagg 21960
atatagaaa gtgtgtgttc tctagatgc ttgcaggtat gaactgctac agccttttag 22020
gagttaatga tgtatgtatg ctgtatgta tgtatttgag acagggtctc gctctgtgc 22080
ccaggctaga tctgttcag tgctgtgatc atggcttact gcagcctga cctcctgagc 22140
tcaatagatt ttcccacctc agcctttcaa gtactgaga ctacaggagt gtgcaatcat 22200
actcagctaa tttttaaat tttttaga catgggggggt ctccaattt tgcccaggct 22260
ggctcgaac tcttgactc aagtatcct cctgcctcaa cctcccaaag tgctgggatt 22320
acctggatga gccactgtgc ccggcctcaa tatctttaa aacagaaatg gacacactct 22380
ttgactagga atgtatccta taaaacact tatacacatg cagagacaca cgagcaagca 22440
tgctttgtaa tagcaatgaa ggctggaaaa actcctcaat caggtaaag ctgtcaagt 22500
cacctgtgta ctatgaaatg gcacttggct ttaacaaga gcaaagacag aaaagcaaaa 22560
gtacaaagta ggggtgtgatg gcacatgcct gcagtcccag ctactcagga ggctgaggca 22620
ggaagatcct ttgagcccag gaggttggagg ccaggagctg ggcaatagtg agaaaaata 22680
aaattaaata ataataataa taaaataggc tgggcacagc ggctcatgcc tgtaatcca 22740

acactttggg aggctgaggt gggaggatcg ctgatccca ggagtcaag gccagcctgg 22800
gcagcaaagc aagacaccca tctcaacgac aaattttaa aaatcagcca ggcaggctgg 22860
gcatggtggc tcacgcctgt aatcccagca ctttgggagg ccgaggcagg cagatcactt 22920
gaggtcagga gttcgagacc agcctggcca acgtggcaaa accctgtctc tactaaaaat 22980
acaaaaatta gctgggcatg gtggcagatg cctgtagtcc cagctactga ggcacaagaa 23040
tcgcttgaac cagggtggca gaagttacag tgagccgaga tcgtgccacc gcactccatc 23100
ctgggcgtga gtgagactcc tgtctcaaaa aaaaaaaaaa aaaaaaaca aggagccagg 23160
cacggtgggg tgaggaggagg cacagaagca gcgcctcttc tgggggcacc cccaatctct 23220
agcgatccag aggcctcagg atcctgaagg gagaaaaaac gtgaagctcc gtgctagaag 23280
agaccataga gattggaatc agctggttct attttacaaa aaaaggaaac tgaggccctc 23340
agaaggtgag tgccttcaa tgcacccacag ggaggcaggg agagggtctt gagccctgca 23400
gggccctgga ttcttgcaat ggggtggagt ggagcctgtg ccgccccac caggcacctt 23460
ctcaggagag gagccgttgt catatccttg aaggggtcct tgagcccctc aaaaggctaa 23520
aaaccacttt cctccttgag tgaaccttca cctcagttta accacaagaa aaactacatt 23580
aaggcccagc gcagtggctc atgtctgtaa tccagcact ttgggaggct gaggtgggtg 23640
gatcgctga gccaggagt tcaagaccag cctgggcaac atagtgaac cctgtctcta 23700
caaaaaacaa caaatcagc tgggcgtggt ggtgcacacc tgagggtcca actacttgcg 23760
ggctgagggtg agaggattgc ttcagcccag gaggtagagg ctgcagtaag cggtgactga 23820
atcactgcac tccagcctca gcaacagagc aagactcaaa aaaaaaaaaa aaagcaggcc 23880
gggtgtggtg gctcacgcct gtaatcccag caccttggga ggccgagcgg gaggatcagg 23940
agatggagac catcctggct aacacgggtga aaccccgtct ctactaaaa tgcaaaaaat 24000
tagccggggcg tggtgccggg tgctgtagt tccagctact caggaggctg aggcaggaga 24060
aaggcgtgac cctgggaggt ggagcttga gtgagctgag atcacaccgc tgactccag 24120

gcttctcatc ccagcacagg tagggggtgc tatgggaaag ggaicctcag ttggccctgt 24360
cactgtctcta tcagctgggg acgtggcatc ctagtgaaaa catcatggcc gggcgcggtg 24420
gctcacgcct ggaatcccag cactttggga ggctgaggag ggtggatcac ttgaggtcag 24480
aagttcgaga ccagcctggt caacatggtg aaacccatct ctactaaaa tacaaaaatt 24540
cgccagggtgt ggtggcgggt acctgtaatc cgagctactc gggagggtga ggcaggagaa 24600
tcgcttgaaac ctgggagggtg gagcttgacg tgagccgaga tcttgccact gcactccagc 24660
ctgggcaaca gagtgaagacg ctgtctcaaa atctcaaaca acaaaacaaa caaaaaacaa 24720
acaaacaaag cgtcatttat ccagcacccc tggggaacca tgctacctgg tgtttatgg 24780
tacctggcaa ggtgcagggtg aagtgtctgc tcttgggcat tgaacccgtc ttgtttgggg 24840
cagctcaggc ccagggcagg gtccgggttg gctctcgttg gtgtggccct ggcccatcca 24900
gacctatatt tctgccgtcc tgcagggtgat caatgttgat gggacgaaga ggcggaccct 24960
cctggaggac aagctcccg acatttcgg gtacacgtg ctgggggact tcatctactg 25020
gactgactgg cagcgccgca gcatcgagcg ggtgcacaag gtcaaggcca gccgggacgt 25080
catcatgac cagctgccc acctgatggg gctcaaagct gtgaatgtgg ccaaggctgt 25140
cggtagtcc ggggggtccc aagccatggc tcagccatgc agacttgcat gaggaggaag 25200
tgacgggtcc atgcctgggc ataagtgtg agctcagggtg ccccgacctg gggaagggca 25260
ggacaggaaa ggtgacagta tctggccaag gacagatggg aagggaacaa gggagctgat 25320
tagggagtgg ttatggacta ggaatgtcgg taacaatggt tagaaagtga ctaacattg 25380
ttgagcacct gctgtgtgcc cggccctggc cgggagcctt cgtgccaca gtgacccgt 25440
ctgcaaagt agttccttgc cctactcgca ctggggagca ggacgcagag ccgtgcaact 25500
cacagggtcc aagctcagga ctccctcctg ggtctgcctg ggctgggctg tgcttgttgc 25560
ccctgtggcc cacgcatgtg cacctccac ctgaaagcca ggatcttcag gacgtcccc 25620
gaggaggtcg ttgtctggca caatgattg tctcttctg aaaaggtgac agagttacac 25680

cctgggcgac agagcaagac tccatctcaa aaaaaaaaaa attaatctc aaaaaaatt 24180
acattaaggc aaactaaaag atgtttaaaa tatatatatt aaattaaata cactccaata 24240
gagcaaatac gaaaataccc agaaaacaca atccccgcac cccagagaca acctcccagg 24300
gggtccacag caagagaccc caagcacgag agacagagaa cagtgtccct gtggcggaac 24360
ctctggccca tcaggctcta ttagaaaata aggccttgc cactgagaga aagaggcaca 24420
gtcgcccagc agccacgggc tctggcacac cacgagtcag gccagcaaag tgtcaactgc 24480
cccctacaag gtgacaaact aggacaaact ggaaccaga ggctggacct ggagcacagg 24540
gaccaccaca tggggctggg gaatgggcag ggacctcaga gcgccacca catgcctaag 24600
agcagcgcgt atgcgcatgc ctctgcatgg cttagggaca caggagctc cccccaccc 24660
caaccagga aggcagcccc cactaccag gtagggaacg gataggacca gcacccggtt 24720
ctgctcgtaa ctagggtc caggccccct cgggggcaac cagcacagag ctacacccc 24780
aaatatcttc acccacctcc tgggtcccat ctggacaagg gtgctgggga ctggctctca 24840
gtcacaccct cggggctactc ttcaaaggac agctggatgc cccagggcag gagcttttgg 24900
ccccagctc cctcacccca gacaccagct ctggggaccc caccagcatg ggcaagggtg 24960
acaccatcgt cccgattttg cagatgagga aactgaggct gagggctggc acacggctct 25020
ccagagctga agagaatgca gagagcagcc ggagccagcc ggtgggtccc tgaggccggc 25080
tcgtagcaag ccacagctgc ctccgcccac cacttggga cctcactggc cccaggacag 25140
ccctccaggg cggcctggca cagagccac accctgctgc ttctgaaca aataagtga 25200
caaggccacc aagccgagga cctggatgta gccccggctc ccgccagggc ctcccaaca 25260
gactcccat ttggagagcg cattaagtgt ttcaaagcc tcacaaacca cagatgtccg 25320
gctgtctcac ggcttctgta acctgaactt ggccctcact ctgccctccc agcactctc 25380
tcagggccca ggccccctct ctgagatgcc agcactgact cccaacttg tccccatcac 25440
ctggctcgtt cctgaacctc ggcaggagag tctcaggcca gatcctcca ccagccacct 25500

ccaccaggat gcaggaggca tgagacctgc tcgtgccggc tgggagatgc aaccaaccaa 25560
gatcaatcca atcagcggat gaactgacaa atataatgtg gtccctccac acaatggaat 25620
attaticagc cacaaaaagg gctgaaatag gccgggcgtg atggctcaca cctgtaatcc 25680
cagcactttg ggaggccgag gccggcagct cacttgaggt caggagtca agaccagcct 25740
ggccaacatg gtgaaatccc gtcttacta aaaatacaaa aattagctgg gcgtggtggc 25800
gggcacctgt aatgcaagct acttgggagc ctgaggcagg agaatcactt aaaccaggga 25860
ggcagaagtt gcagtgagcc aagatgcac caccgcactc caacctgggc aacagagcaa 25920
gactccattt caaaaaaaaa ataaaaggct gaaacacca tacgtgttac tacttggatg 25980
actcctgaaa acgttacagt aaccaaggaa gtcagccacg aagacgcatt gtaagattcc 26040
cttcatgcaa aatgccaga acaggcagaa ccacagaggc agaaagtcga ctggtgttca 26100
ccaggggatc cggggagagg gaacgggaag tcaccgtgta atgggtatgg gttttattt 26160
ggggtgatgg aaatctctta taactigata gaagagaggg ttgtaaacac tgtgaatga 26220
ccaaatgcct gccttctata cttaatat ttatattata taagtttcac ctcaatttaa 26280
aaaaaaaaa actcgacacc ttacacatg gaaagatctg gctttagctt gcatttcctg 26340
taactcctgc ctaaagcctt ccagaagctt ccgtgcctt gtggatcaca accagactcc 26400
acaccatgat ctggcctcta agggcctctc gcaggacacc ccgagggtga aggagcacc 26460
gtgggcccac ctctgcatag ctgcaaagct tcttccctg tctccctc tacatgggaa 26520
gctctgcccg caggggcggg gccttatctg ccattctatc gcactcaacc ctagcactc 26580
actcggtagc agacaccaa gcaaacagc aacagcatta taccgggcca ggtgcacgtt 26640
aactcactga attcatggta ggaaggattc tattccatt ttacaggta gaaaactgag 26700
gcacacaaag gtagcatcag cticctaagc ctcccagcac aggaagcggc caggctggaa 26760
tcagaccctg ggcgcagggg ctctgtccac agtgctaact aactactct gccccgagg 26820
gctgcagcgg tgagtgagt agttgtcag tggactggat gtccaaggctc atacaggaaa 26880

aatccagact attgtaataa cagcctctag accggctggg gccagaaaga tcgaggacgc 26940
tgacacacaa ctgcgctcac tgcagctctg ccagggatgg ggctaaaggt ctacacagg 27000
gcagttaggg ctccccatag cctgggagag gaacgggggtg agataacaga aactaggtat 27060
ggtgcccga gtaaacagc cactgagcat gtaaaccag gtgggtctga ccccaaacc 27120
ctccaccccc atcagccctg caaccgctg ctgcaaggga gaaagcaact cagaggcctc 27180
acctgcctac atccccacc cgtgtgtgtg agtttacta aatgcctgag cagtgcaca 27240
gcacggctga aattaacgg gticcaaaaa cgacaggaag cacgaagtga atctcccag 27300
gaaagtgtg aacaaatgt ggatcgggtt caccggcgaa ttcttgga ctgaagagg 27360
gagctaaaca cacggggccc tgcttgagg gggactctt cagggtgctc cacacagc 27420
ttggttaacc cactcagcc ctctgggct ctccagagg gcccggcctt ggccttgggc 27480
atctacagga ggaacctca gggggagagg ggggtgcctgg acaggccggc cctggaaca 27540
gcacttgggc cccgaggaga gaggactagg gcttgggagc tggggaagti ctgacactg 27600
ggaccactag aacaaagcca ttccgtgcg ttcacagctt ccaattgca caggaagca 27660
tcaggaaaa taattagcgg ccacttact ggctcgctg aggtccgagg catgtattc 27720
acacagtaaa accagggata taacatcaa accgttctgc agaaagattc ctcccttcc 27780
ttccattta ggcctggatc accacattca ctggggctcc caggccttgc tgcctaagt 27840
taaaataatc aactctatt ttgctcaca cacaactgaa ctctacagct ataattctt 27900
ctcctcagg gtcgaacca catggacgac aggcatttga ctccagcaac atcacccaa 27960
aacgtgcaca aaacccaaaa ctgcaatgag gtgaaaggca acgcggctcg cctagaaacc 28020
cccccttaa aacaaacagt ttcccaaaa cccctttgc ctcttgacc caggcattc 28080
cggaaaaagg agcggcgctg gcctgtactc ccagatact gtcgtgttt tgcctcacc 28140
ttgtttgct agctccagac aaggccccc aatgtaaca cgctcctgaa agaggcagat 28200
ttggggtgaa actgtcata gaatctctag gcttgggtca gaggcaggag gacgtgaaac 28260

aaactccaag ctctctctgt tccccgtgt cccccacacc tccaagcaga ggctgcagcc 28320
tgggggatct gactacaggg ccaccccgct gcaccattca cactggaaat attcaggag 28380
acagctgttt gccttaagga ggcccagaca aagggggccc aggtcctccc cgctaaactg 28440
ccacaaacag aacaggagcc gcggcgtgca caggcacttg cggccgtgcc acttggccag 28500
ccatactcca gaaaaacaaa acacgcacat ccgaagagaa tgatttaggt agcaagaggc 28560
ttgcttgaag aaccacatgg caatctccaa attaaaagaa catgtgtagc gtttcacgac 28620
tgcttaagtt tcctgagtc tcctgacctc aactccaccc cctgggaaac accaaaagtt 28680
ggatgagaaa gtccccccgc cctacctctc cccacgggag tgtacaactg aggcacaagc 28740
ctgcctcccc cactgccccg cgatctggga ccacgtctcc tccgcgtagc cgacccgggg 28800
atggacacta tctggggacc cggcggccac acggggcatt cgggtcgccc gggcacctgg 28860
caggtgtcag tccgcttga aaccacagc cacgcggctc acaggagcag cgccaccggc 28920
taggccgccc cgcgcccggg ctcagaactt tctcgtgcc acttcagccc gtctcggag 28980
cacgcggggc ggccgcgcgg ccgctggaaa caggcttgcg aaccggctcc ccgggccagg 29040
cccgcctcgc cgccccaagt cccgctcgg tgcccgccc gggccacacg ggcccagcgc 29100
gggctcggct cggctcccgc ctcccgcgg gctcgggcag gtgaggacct gccgcgccg 29160
cacctggcgg agcgggcgcc ctctcgcca gcccgggacg cagcgtcccc ggggagggcc 29220
cgggtgggga gacaaagggc ccgcgcgtgg cggggacgcc ggggacggca gggggatccc 29280
gggcgcgcgc cccaactcgc tccaactcg ccaagtcgt tccgagacgg cggcggcgcc 29340
cgcgcacttg gccgcggggc cgccggggcc attgtccgag caaccgcgg cccgtctac 29400
acgccgggcg cgggaaggtc tcgaatcagg 29430

<210> 8

<211> 33769

<212> DNA

<213> Homo sapiens

<220>

<221> unsure

<222> (33739),(33749),(33758)

<223> Identity of nucleotide sequences at the above locations are unknown.

<400> 8

```
ctcccccta cactggtcct tcgacccgcc tcggatgaaa actgaatggg ttagccta 60
gaggctctcg gtctctaagg gaggtgggtc aggatgccgg ggacagggtc ctcttctgg 120
ggcaacgtgg gggaacgagc cacctacccc tccactgaat tgccctgggg tgtgggtacc 180
gacggctcat tcggtgtcca gggtctgaga tgtgtgaca ggaagaatga aaggggatgg 240
gagggatggg gcgaaagaag ccacctgcag cccaggaac tatctggcca gcacaccgtc 300
accagcggc ctgagccacc cctgccagag ccaggaggag accctgcca tgggtcacca 360
gtgtgcagga actcagaagg tcatcacagt taataccctc catgccccaa tgtgggaaaa 420
caggttttt cacaacaaac aagataattt ttgtatttt ggcaaaagga ggcagggcag 480
ccccggacac ctccatcca cctcatcacc cagccgcagg gccccggcca tccctgcaga 540
cagagtggat gtcacaacct ccctgcaccg aaccaagtgc agctcccagg ccacaggcca 600
cccaggaaag gtccagtggc ccccgagggc tccaccgca ggctcccac cacagccggc 660
accaaccag gatagctgtg ttctctggc ttctttcac acgggtagca gaaagctgag 720
atccggggaa agctgagatc cagggaagc tgagaatcgg cctctgctgc ccggacgccc 780
acccccagct ctgctccag ctccagggcc tccttctcag gtgcccttac aggaggcaga 840
```

gggcttgagc cacctcctgg gcctggggca cgcaggatga acggggtcac ggtgcaggcc 900
actgtccact gcgcagatcc caaggccata aacagcctgg ccacagtggc ttcccagctg 960
gcaggcggcc agattatitt tgtgttttag caattgatta agttttccg ctgccccag 1020
gggtaagtgg tggggcaaatt gccgcaaccg cagcattga cccgggatcc tgtccaagt 1080
gaccataggg tcacaaagca caagggaagt ggctggggcc gatgctggct ctgctggaac 1140
ctgaggccgg cactgtcac ctgcacggtg cctgggacct tccagcaagc acagagaagc 1200
tatggccctc caggagcagc tggcaggcac ctggcctgc agtcaggggc tctgtctgct 1260
cagctctaaa acaggaaagt cgctgctctg cctgggggtca gggcagccag agagtgaac 1320
agtcagtgcc ggccctagga agggacctgc aggggggtcc ctctctctcc catccctcgg 1380
tgccagccag cccctcctgt ggccccccac tgctgcctc tgccccatg ccccaccaca 1440
acctcaggcc catggctgca tggccactcc ccaggcaggc agtggggatg ggatttcacc 1500
atgttgcca ggctggtctc gaactcctga cctcaggta ggagttccta aagtgtggg 1560
attacaggcg tgagccaccg cgccagccct ccctgtgta ctaaactc acacccctt 1620
gctggggacc ctggtgagg aacacagcct cacaagtga gtgtggttt gtgagcaaa 1680
tgacgcctgg gcagccctct catcttgcc taaaactgaa gaatttaggg gcgtggatgt 1740
ataaaacagt tggtagctta aatgaaaag aaggccacac tcccccttt aggcaggcgg 1800
cctaattctt taaaagccag cacagggtgc ctcttgaac ccaggcacac agtaggtgt 1860
caatggacag cagcgggtac ttgtactgt catgacacc tgtctgtggc ctctgcagct 1920
ggctccagcc tgacgcatgg ctgcgccct ccgcaaggcc accccggtat acatggaaac 1980
tctgtggaga aggccttggg ggccggccag gacgccaggc ccagatcca tctgcgcct 2040
tctccatag acctcagcga gctctggca ccatgtgcct caggccatt taagaagtag 2100
ggccggccag gcatggtggc tcatgcctgt aatcccagca ctttgggagg cccaaggtgg 2160
gtggatcacg agatggtcag gagatcgaga ccatcctggc taacacggtg aaacccatc 2220

tctactaaaa atacaaaaaa taagccgagt gtggtggcgg gtgcctatag tccaagctac 2280
tcgggaggct gaggcaggat aatcgcttga gctcagcagg cagaggttgc agttagcgga 2340
gatcgcgcca ttgcactcca gcctagggtga cagagagaga ctctgtctca attaaaaaaa 2400
aaaaaataa aaaaaagaag cagggccagc cacggacgac ccctcacaca gctcccagga 2460
cgcggtcctg ggtatagggc tcaggacat gaccgctgca gtggcccca agaaacgtta 2520
ctttgtcac ccaccccgcc tcagtggcag tagccaaaat aacggattag aatggaacca 2580
tgtgacaatg ccactgcccc aactgacaga agatggctat cagcagtta cgcgggccca 2640
cctatcaca gtgcaggga ctctacaact tatgcatcct tcccagaca ccgtccttc 2700
gacctcca ggtcagcaag gcacacagg cctacattc acagccacac agcagaggc 2760
tgaggctga actcggatgc tctgatttc gttcaatcac atcccagag gtggcacaga 2820
gacggggggc ttctctgac aaagtcaaga aagtcactgc cagctccact gaagacaaa 2880
gaacctcagc tctaaacc tctgaagg gttaccgaac tctccagcc tgttcctgg 2940
gtcccgatg tggccccgtg ggacacagga agaggaagaa gtccttaga gcagagcctg 3000
gtgcacctgc cacactctca gagggctgcg cacgggcgga ggagccgtgt gcaggagtgg 3060
ggtctgatg gaggggcgct gtggccgggg gcagggggca ggggaagggt gtcaggtg 3120
gtgggcacag cagagcagg ggcaggagg tccactca gatgtgcaca gggagaaca 3180
aatcgtcat ttccattgga ataggcggt aaaggtaga aaacagagt ggggccagga 3240
agggagtcgg agccttctag tgtctctg caggtgagc gcagccgag gtgtcagctc 3300
agcagactg ggtccagg ggcgtgtct ctactga cccaggga cacggaactg 3360
gggagggaga gcagaggcac agggcacggt cagtgaacg aaacaaggag tcatcacia 3420
atcggaag ggcaaggagt gcccgcagc gcacaagggt tctgtctgg caacgtggg 3480
gtcccaccag gccccgacc ctgcaagcgc aaagctcgc actgaagata aagggaagct 3540
gttgagctg cgagctggt ctgggtccg catggagctg ggcttatgt gcagtcaca 3600

gggggacatg gaagaggctg caggggacaa aaccagtgc cacagtctaa ctctgagcct 3660
gtggaaaggc gccacagca ttcacccatc ccagagatgc caticccct gtgccccgc 3720
tccacggtga cagcgttctc caggaatatg atgcgccct ctctcttgc atcagccctg 3780
acagtgagta ttcaggccaa aaagcagaag agcacagctg cgtggttcca ttccatgta 3840
gttctggaac aggcaacgct aatccaaggt gatagaagtc aggagagtgg tggagggggc 3900
gggggttgag gatggcaaag gggcaccggg aactttcca gtggtagaaa tgttctctgt 3960
ctggaccgtg tgtagttat gcagacatat gcagctgtca aagtaatcc aaatgtacac 4020
gttaaatgt gtgcgttta ttgcctgcaa gtatacctc aattaataaa ataaagttag 4080
cactcaggct tctccacaa ctctctgaac cgtgtgagct gatttcttg ctattaaaaa 4140
ttcacggtcc atggctgaga acagcagctg cttctgttt gcaaagtcaa cgccaatcac 4200
tgcccgccg cggcagactc ggccccacag gacctcttt cttttccc ttgacctac 4260
ttccctgata agtgacaaga cagccagact ctgggaacaa acgcccgtta ttcggccccg 4320
agctgagcgg gccctgcttc ctgagctaata ccgcccggac agacggaggg acgtgagggg 4380
cttgccgtc ggctccagct gtcagtctgc ccgtcagact cgacagtggc cccctctgtt 4440
cctcccgtg cccccactcc atccccgact tcttttgtt tctgtccct gacagacgaa 4500
catctgttaa aactctgtct gggtagctg tggccagcgg cccacaaatc cccaagccgc 4560
acccagcct catctgggcg ctgccgggag cactgcctgg ccaccctctg gacatagctc 4620
tgagagccac cggccagggc acgtgtggcc cgagtggcat ggtgcacgcc gctaagccca 4680
ctgccccaaag gccccaaagc aggagggatg tgcaggagac aaaagtcaaa agaacagggg 4740
cacgttcac agaggatggg gctggagggg tggcagttag gaacagcagc ttccaggat 4800
ggcgggtggca actcccaaat aaggcctcac tctgtctgtt ttagctcat tccacataat 4860
tgaaaaaca tggcagaaac cgaagccagc tgctgccttg gtctggggc tgtgtggagg 4920
gggtggggag gccggaggcc caggctctgc actcgactgc tgggatgag agtgactctg 4980

agctgcagag agcagcatcg cagccgccat ggtccattg agccccggcc acgctgggcg 5040
gcagaggctc gtgggatata cctgccctgt ctcatggggg tcacttcagg aggggcgggg 5100
gagccaggac acagcccagg gctagcggtc accctgcagc tcaggggcca cgtaaatagt 5160
gccacctga aggcacacag cagtgcgggg cccccccgc caccaacgca tccctacctc 5220
taggaggccg cctgtgtgcc cctgggaacg ctgctccctg tcccttgggg tctgtgtgtg 5280
accacctct cagccccctc ctgggggaag gcacctgact ccctacacc agctggcttt 5340
catttctca aaatcaggaa aaagcagaat tcaagacatc acagaaatgt ctgcctgt 5400
aactccatga aagataaacg gtcagacacc caggaggag tccaggagc ccttgagtct 5460
cacctgaggc tctggctca aacctgaga tgtttcagc catgctagcg ccgccccca 5520
caacctgccc cacacagtcc tcccttgga actcacagat ttggccccca cctgccccgt 5580
ttcttctgt ggagtgggtg cgttgggtg ggggtgggct ggggactctg gatgtgtctt 5640
aagagtctga gtgattctga cacagccagg ccctgcccc ctcctgacct tcgccccaca 5700
ggaaaggag ccacacgcct gaagcgcca gcacacccc ctccgtctc ccaggtcac 5760
ccgctggccg tgtgagccgt gctccccact gccccttcac ccacccagc tctcctggc 5820
agcaccagc ctggaagct acttctgatt acaaccgcg aaggaagact cgtccctcg 5880
gcactgacc agacagcctg caccatcac ctgctcagca caaccacac agccttcctc 5940
caaacccat ggagcgggga gtataatcac ccccttcta ccaacggaca aactgaagca 6000
cagagagggt aagtcactt cctaagctc caacacgatg aaaaaata gaaggtcagc 6060
ccgcaagtgg aactaggtgc tccaagtccc cggctgcct gacctgcac ctctcgccg 6120
ccacgggtccc gggtcgcct gacctgcac ctctcgccg ccacgggtccc gggtcgcct 6180
gacctgcac ctctcgccg ccacgggtccc gggtcgcct gacctgcac ctctcgccg 6240
ccacgggtccc gggtcgcct gacctgcac ctctcgccg ccacgggtccc gggtcgcct 6300
gacctgcac ctctcgccg ccacgggtccc gggtcgcct gacctgcac ctctcgccg 6360

ccacgggtccc gggctctgcct gacactgcac ctctcaaca ccaccacggg cccgggtctg 6420
cctgacactg cacctctca ccaccaccac agtcccgggt ctgcctgaca ctgcatttcc 6480
tcataccac agtcccgggt ctgcctgaca ctgcatttcc tcataccac ggtcccgggt 6540
ctgcctgaca ctgcacctcc tcaccgccac ggtcccgggt ctgcctgaca ctgcacttcc 6600
tcaacaccac tccttgccg gctcccaact acaaaccaag ccatgtcttc catcctgaat 6660
cctcttgcc taaacatcac tcacaatgcc tccctcgga acaggcaca gtcccaccag 6720
cacagcctcc ttcgttacct gcgttccgc tagcccagg ccagctccag agccctcacc 6780
acagagcctc tctcttcac ccccgacac tggacctcac caaccatag cctggaggag 6840
atccctgtgt gacccagggt cctcctctgc ccgactctga atttactgc ccaacgtgac 6900
acctcggaag gctctctggg cactggcagc cctcatggg caccgtcct tctggccagc 6960
tctgacatcc cggctgggtga ggtgccctgc acgaggcctc tgcccactgg gacctcacag 7020
ccgtgtgtc agctgaaca agcgacagaa ttacagtt tcttcagtt gccctgggt 7080
gagcagctcc aggtagttt cagtcgagge gaggcgtccc gtcagcagcc aggcggcaca 7140
gctaattcat gcccgcggg cgcacggccg caatacaat gggcacctgc agcctggaaa 7200
gccacagagg aaccgagaac agcgactgtg ctgaggtgac aggactgtgg tctttaaca 7260
aaacatttc cttaacgtg atattttacg gcaaggaatg aaacctggag ggcaggacat 7320
ttggatacta aagccccagg ctgccgctg gtctgcttg tgaagtctga agcccgcgcc 7380
ccattctggc cccgtcaca ggtccggctc tgactacca gcttcaatgc taggccgtgc 7440
ctgtctcca accagaacat gacttccta aggacaaagc cgttctcgc ccatcccat 7500
ctcctctgg attaagaaat atgggaagat ctctagaac cacctcaat ttgcagagag 7560
ccatcttgt gacaaacct tgaaatgctt ctaagaagag tttaggttc ttctcaactc 7620
taaaacctct agaaaactct attccacac cagctgcccc tggaacactt cagcttcaaa 7680
agggcccagg gcaggagac ggaggagcca gcatccacac cgagcaccag cctgttaatt 7740

aacgggaagc ggggtggggcc catctccagg cagctctgag gtcagactgg ggaaccatgc 7800
ttacaaaaaa aagtgaactg aaacgctcac gtcctcatgc aaaaccagac tcccagttgc 7860
atctttctgt ctcatgagg agctttttcc tccctttgac agaaccacct acacacggca 7920
tctggaacca aagcagaaag attcaggctc agagtaaaac agtccccaca ctggctgcat 7980
gtggacgttc ccggcccaga gtctcgccca agcagggcct ataaatgaca caaatgttt 8040
ttctctgcg tgccagtcac gctccaactg agttatgtgt aaaagtcct ctcacggctg 8100
agggcaaaaa cagttcccac aagactagag aaaggtagc cctgacggct gagtctctag 8160
ggagcgtgga gctgcgtgct cagccctgcg gccctgacgg ctctggaatg gaaaagctat 8220
ccaactggaa gggcagggct cgtgctagt ccagcggctc aacccacag gtgtctgtgg 8280
gtcagctcc atgccacaga gcccagggt ggggccagag ccaccaggcc ccctgccagc 8340
ctgcaggggc ctctcctct gggtagccta accacccct gtgagcgcag gcagcctcct 8400
ctaatacca cagggcctgt cccccctct ccccgctg caggaaaatg agccctgagg 8460
actccccagg gctgctctgg gcctggacat ggagactggg aattacatt gcagaaggag 8520
cgcaatgcc tgaagggt cagccacgag cagccagtcc ccagggtca gaaggcccag 8580
ctgttagaac cctgggagcc agcaaagagc caggggtcc acctaagtct atagcccctg 8640
cctcttctgg ttgggaaaga aatcaacgcc ctttactgg ctccactga cagcccactc 8700
ccccaggtat gggaggattc tgggacgat caggcaaacc tggaccctga gtgaacctgc 8760
cccagctctc acgggccttg caccagccac agcacctaag gcgccgtca tggtgacaac 8820
atgaaggtga taaggcatg gacagtggac atggcagctg gacactgggc acccactgga 8880
tgccaggcac ccagcacggc tccgtaccc ctggatgagc agtggccctt tgcaagccag 8940
ggtagcctgg gcaagttatt tgggggtctc caagctgtc cagctgtgcg acttactga 9000
gccatgagtc tgggatttta tcagggccca caccgttc tggaaactctg atacgtgagg 9060
gagccacaca gggaccctta aaaaagctc ccagggaac atgttctt gcctcagtct 9120

cccaaatagc tgggattaca ggcgacagac taccgcccgg ctaattttg tatttttagt 9180
agagacaggg ttccacatg ttggccaggc tggcttgaa cccctgacct caaatgatcc 9240
ttccactgtt agggcaaggc acctgacagg cacgactgca cgatctgctt gttgggggct 9300
gtgtccattc cccactcctt cgacaaatgt ccacaccag ccttgctttg acacccaag 9360
aacagagatg gtgacacctg ctctctacat gccattgct ctccaaggc agacatcccc 9420
agcagatgca acacagtgt taggcagaca tcaccaatcg atggtggcaa cagacaccag 9480
gccctgctcc ctctaactcc agtggccagg cccaagcca gctctcacct gccactccc 9540
aaccacagc agcaagactc agaaatggca aaaacacaaa gagaacagaa acgccccata 9600
gcgaggaggat gactaaaaga catgtcttga taagatattg ttcaggcata ggccaggcac 9660
agtggctcat gcctgtgatc ctagaacttt aggaggctga ggtaggtgga tcacctgagg 9720
ttaggagttc aagaccagcc tagccaacat ggtgaaacct catctctact aaacatacaa 9780
aaattagcca gacatagtag cgggcgcctg taatcccagc tgcttgggag gctgaggcag 9840
gagaattgct tgaacctggg aggtggaagc tgctgtgagc cactgtactc caacctggac 9900
aacagagcaa gactctgtct caaaaaaaaaa aaaaaaaaaa gatatccttc actaaaactc 9960
atgtcttga tacatattta cctctgcaa tcgcaaatgc ttctgcagtg cataaagtga 10020
aataaatagc aggaagcctt acggttcgat caccacaca gacacacagt cacatacagg 10080
aaaaacgcag ggagggtgg ggaacaaaaa aacagaagat aaaatgtgga gacagacaca 10140
ccaagagagt aagagaccac ctccagacct ccttcagct tctaaacac acgagccggg 10200
cccgttacag aatttgcggg gaccgtgca aaatggaagt gcagacagcc ccttactcaa 10260
aaggtaggaa tticaggtca acaacagagc tcacctcata tgactacaca ggtcacacag 10320
cccgtgaagt cgggtccaac accagcatgc tctgcctca aagccgtgc acgtgctgtt 10380
ccttctgcc ttccctctt ttagtcttc agatctcagg cctctgaga gagaccttg 10440
acctgccggc tcaggcggcc acaccccag tacaggagtc tccggctcag cccctgctgt 10500

gttccgtacc cgatccaggt ctgtcctatg tccatctgtg tgccggcttg ctctctgaca 10560
tgccccccac cacacgtgtg cctcggggca ggggaacagg cccgtctcat taactgctt 10620
cttctcagat attttctgga atattgtgg atattgggca acatatatgc tccaccttt 10680
tcagactagc caggacgagc tgcattttt tttttttt tttgagacag ggtctcactc 10740
tgttgcccag gctggagtat agcggcatga tcttggtca gtgcaacctc cgctcctag 10800
gctcaagcaa ttctctgcc tcagtctccc aagtagctgg gattacaggc ccgtgccact 10860
actgccagc taattttat attttagta gagatggagt ttcacatgt tggccaggct 10920
ggcttgaac tctgacctc aaatgatcca cctgccttg actcccaa atgttgggatt 10980
acaggcgtga gccactgcgc ccggcccgag ctgcctgtt tacaccttg ccatattccg 11040
gtgattctct ctcccctcg tccccggcc ctgactgtg tggccactcc ctgccgtcat 11100
gagcccgat gtctcactc ttcccttc cgccaggact tcaaccaaca ctgcagagcg 11160
cagggtccag ctccagcact gattcagcc tctctcacc aacagacagg caggaaagaa 11220
aaciaactct gagaaggcca aggttcccg gcagccagca agccaagcat cttctccgc 11280
tgaggctgt gcagccgagg caccctcc tccaggagc aggcagcgtc ctggggcagt 11340
ctgcgaggga gaccagggcc ctgtctccac cagggcccca ggtatggggg cagcagcaa 11400
ctcatggctc tgggagccag accccacctg ctagaacctc ctatgccacc tgctgtgggc 11460
aaccagcgc tggtgacttg ccctggcctc ctctgtaaac aaagggtca tccaacctgg 11520
tcaaaccact cctcccctc aagggtctat aatctcct taacctgctt ggtccaaacc 11580
cctggtgtcg ccaggtcact caggaggcag ctcatctgga ctcttcctt ggtccagt 11640
tctctctcaa cattgcctt gaggccgagg tgaacggtca acagcgaagg gcccagagg 11700
tgatggagga gcgggtgtcc aagacactca cctttctaa tgcactgact ccctcgtgga 11760
ctcactgtg ccgtctccc caccaccca gcccagagc ccagagtgcg agcgcagag 11820
gcccgggatt ctgtctgcac cgcggggtcc ccagtgcctc ggagcaatgc cagcaccgg 11880

caagtgttcg acaaatgcct gctgaatgag caaatggatg gatgaacgaa tgaatgagca 11940
agcagatgaa tgaatggggt gctgtccaga gccgtgagga ctaggccgcc caagtcccca 12000
tttctcaaat tctccttctc ccgacttggg aaacaagatg cttggtcggg gaggctctcc 12060
aaccatcccc tgcagcagcc ggcacagcgg acagaccctt tgatgtaaca gccatgtctt 12120
cattaaagat gccctgctct cagaaagaga aagacaaata caaacctgga aaatcctcac 12180
caaacgcagg acccctgcca gggagcagag aaaagaccca cagccacgg gcgccacgac 12240
cacacacaca cccagccgc tgcacacaaa cacagaccct agccagcaag aacaggggga 12300
ccaggaaact gttcctaaag tcaggacccc catgtgtca gacagcagt agagcaagga 12360
cacttctcca tccaccgat gccaggagag tcttttagg gggccccaca ccgagactct 12420
gcccttagga ctgttctga gtgtggaagc cagcccactt ggaagcccc tgcctccccg 12480
agtgggacac cggcacagga agcaggccct gtccccacc actttctgca agctgggccc 12540
catcacgcta cagaaacggg gaggactggt cccagggatg gcgcttctt gacacctctc 12600
gttaccctt cgcttgccag gcccagggt cagccccaga ggccagactg gctatcccag 12660
gcccgggagc atccccgaag gcgagctgca tctgaacgt gtgtgattc ccgaagggcc 12720
cgccccgaac cgacacctgg aaagaaagat cctcagccgg tgccccagag gagaagagcc 12780
atgcctcact gcaacacagt cccaggaagc accaagtgcc tgaggaccaa ggcggagagt 12840
aaaaaagtgg aaaatatctg gggcaaaaat aaaacaaaac aaaacaggat tgacctctg 12900
ggctcaagca atcctcccaa ctgagcttcc cgagtagctg ggaccacaga cttgaatcac 12960
cacaccgcc aagtggatca tttcgaacgg gtttccgag gttcctctg gggcaccccc 13020
ggcgcccgca acccattccc gccaggcccc gcccccccg cccgccccgt cccgtccac 13080
cgctcacct gccttacag tctgcccgt gtctgcagc tgcacaccg tggggcaggc 13140
gcatgtgtag aaaggctcgc ttggggacag caggcacagg tgggagcagc cgccattgtc 13200
ctcctcacag cgagtgtgga ctgagaaaac caggacagac tgagagaagg ttccagaaga 13260

ggaccgtcac ttgtttctga atgagtcaca tctgcctcg tccccctga cagcctccag 13320
tgtgtccctc tgcccaaaca tcggcctcaa gtggcatcag ggacctcccc gcggggcacca 13380
ttccacctgc ctcatcgctg gccccgtcca catggggccc tcagcctggc cagacggcct 13440
gcaatttccc caaaaccagc cgtgaccttc ctggccaccc tcacaccag atgtgacctg 13500
cccatggagt gacatcctcc ccattctgctt cctcccacca agctcctatg actagaacac 13560
cctccccagc tctcggagc ccccaaagga caccctctg caaaggctgc cccccagct 13620
ccaatggccg gggtcaggac ctgcctgtgt ggtagtgacg ggaaccccag agacaatggg 13680
ctcctgggca aaaggcttgt cttgtctttg tgctatgtgt ggaccagca gttccatag 13740
gaacactgtc ctcttgctg ggatggccaa gcttgctact ctccaagcc ctctatgac 13800
caacagcaat tgaacggaac tcgataaatg ctccagcac ctattcaaa ccaggggaaa 13860
gctgggtgta gcagcccaa aatacggata taactggaac acaaaactca tcaaatgaa 13920
cctcctctc cctcatgctg cccaagtgt agatgggttt tgtgaccag actttctac 13980
caggaaacag ctccagagag cccaccctc ctgtgtctg ctctgggaac agctggcacc 14040
cctaggcccc acatttcaat tcaaagtcca aacctccat aatggcctgg ccagaaatct 14100
ccatccctgg tccctgtggg agtgggccac tgtccccaga gccgcagccc cactgtcaca 14160
gaagctggtg catttcccca tcaggacact ctgtcacaac ccagcgtggc cccaggctg 14220
agaactgctg attctgggca gattattcat tgataaatc gcgacttga gggccaagca 14280
tggtggctca tacctgtgac ccagcactt tgggaagtca aggtgtgagg atcactggag 14340
cccacgagtt tgagacaagc ctgggcaacg tggcaaaatc tctatctct attaaaaata 14400
catacacaca cacacacaca cacacacaca cacatatata tgtatatata aataaccata 14460
tatatatata cacacatacg tgtatgtgta tataaatata tatacacaca cacacagaca 14520
acttctctg ggccttgaac acgaggcaac ctctcttga aatcccctt cactgtctga 14580
gctgaaata gccccatga gctctgcaga ggggtcctct gcaggcccgt gtccccagc 14640

cagccacaca cctccctcca ttgcagcagg tacccttta gagagggggc ccccagagc 14700
atgggcttct gcaggggagg gtcacctgcc cccccacccc acccagccc gcgcaccccc 14760
acgccccgc atcctccac tccctgccc cgcgccccg cccccccag cccctcacc 14820
ctctccccg tgcccaacc ggcactcaca aaaaggctgc cgctctggc tcagcacctg 14880
gatgtcatg ggtgagtata gggcactcag gatctcttc ctctcccc cagtgcgctt 14940
gttgaggca tggatggagc gggctgcc gtctgtccag tacagagtgt ccccgagag 15000
cgtcaggcg aagggtgcg tcaggctgcc ctccaccacc ttctgctgc agtcaggga 15060
gcggggtgga ggagccatca ggagggtccc ccgacagtca ttgtgtga cccaattaat 15120
ttcttttt tttttgaga tggagtctg gtctgtgcc caggctggag tgcagtgatg 15180
taatctcagc tactgcaac ctccgctcc cgggtcaag caattatct gcctcagct 15240
cccagtagc tgggatcact gatgccacc actagccca gatgatttt gtatttttag 15300
tagagacagg gttcatcat gtggcaagg ctggtctga actctgacc tcaggtgatc 15360
caccacctc agcctctcaa agcgtggga ttacaggcgt gcgccaccat gccaggcttc 15420
ccatttgct tcaaccagac aagtaggcc aggtcaagag cccaggagc tggcgccctc 15480
gtacatttct ccggcgctgc acagggcacc tccaaacac agcctgtgat ggtgacacac 15540
gggctcccc aggtcaagt gcaaagtct cccaggga gaaaggagga agccatgcct 15600
ggcaaaaagc acacctctcc tgccaacgc ttaacctct gtatacaat caggccatgt 15660
gcactcgtc ctcttaca tgctcataat ttactttc agagtaaag aaacttgga 15720
tcaaccgag aaacagctat tctttctag atgttacag tgccagcaa atgaggactc 15780
gggtgtaag agattatgga cactggaaac aggatcataa tgtgacgtgg tcggtaatgt 15840
gcagtttat ttgttaag accctgccc cgtgacagg tccctgagg tgggctggg 15900
ggcagaggc cccgccagc cccagccct cagcacagtt gccaggagag ggtgacactc 15960
atgaagtggc acagggaaga tgggagctgt gggctctga gatccaccac ctcttctgtt 16020

catttttgtt gatgctgttt ttttaagaaaa ttattgaagt aaaattcaca ggacatacgt 16080
ttactttttt tttttttt ggagatgggg tctcactctg tcaccaggt tggagtgcag 16140
tgggtgatc tcagctcact gcaacctctg cctcccaggt tcaagcgatt ctcccacctc 16200
cgctccaga gtagctggga ccacaggcgt gcaccaccac acccagctaa ttttggggg 16260
gtatctttt ggtagagaca gggtttcgcc atgttgcca aggctggtct tgaagccctg 16320
agctcaggcg atccaccgc ctggcctct caaagtctg ggattacagg cataagccac 16380
tgcaccagc ctaaattac cactttaag tgaatagtgt tacctagtgc atcgcaagg 16440
cggtgcagcc tccactctg tctagtcca aagcattcc attgcccac aggcaaacc 16500
cacaccggc agcagtcag cccagctcc cgccccagc cccggcaaac actttgatg 16560
gacttaacta cacacattct caacatctca tataaacgga atcacaatat acagcctctg 16620
atgtctgtct tcttgactt ggcacatgt ttctgaggt catccaggct gtagcatgtc 16680
agtgctcat cccgttttag gggatgaacca tattccagtg tgcagacaga aaccaatctg 16740
tgcattcatt caccactgg gggaccttg tgcattcc accctcggct gttgtgcaca 16800
gtgctgtac ggacattact gtccattcac atttgtgtg aagacctgt ttgattctt 16860
aagagtatac agctaggagc ggaattgctg ggtcacatgt aaatcaatgt ttacgtctca 16920
aggaatcaac aaactgttt ccacaatgt gtctttttg ttgtttct gagacagggt 16980
cttgctctgt caccaggct ggagtgcggt ggtgtgatca tggctcactg cagcctcaat 17040
ctcctaagct caatccatcc tctgcctca gcctcctgag tagctgggaa cacaggtatg 17100
taccaccatg gccagctaat ttctaattt tttttttt tgttttgtt ttttgagac 17160
agagtctcgc tctgtcggc aggctggagt gcagtgggtc catctcagct cactgcaagc 17220
tctgcctccc gggtcacac cattctcctg cctcagcctc ccgagtggct gggactatag 17280
tcaccggcca ccagcctgg ctaattttt tgtatttta gtagagatgg ggttcaccg 17340
tgttaccag gatggctcg atctcctaac ttcatgatcc acctgccttg gcctccaaa 17400

gttctgggat tacaggcgtg agccaccacg cccgacctta cttttaattt tttaatttta 17460
ttattttatt ttattttttt ttttttgag acagagtctc gctctgtagc ccaggctgga 17520
gtgcagtggc gggatctcag ctactgcaa gctccacctc ccaggttcac gccattctcc 17580
tgccctagcc tcccgagtag ctgggactac aggtgcccac cacgatgcc ggctaatttt 17640
ttgtatttt agtagagaca gggtttact gtgttagcca ggatgatctc aatctcctga 17700
cctcgtgatc cgcccgtctc agcctcccaa agtgctggga ttacaggcgt gagccaccgc 17760
gccagcctt tttttttt tttttttt tttagata gagtcttgc ctgtcgccca 17820
ggctggagtg cagtggcggg atctcagctc actgcaagct ccgcctcca ggttcacgcc 17880
attctcctgc ctacgcctcc cgagtagctg ggactacagg caccaccac cacacctggc 17940
taatgtttg tatttttagt agagacgagg ttccacctg ttagccagga tggctcgtat 18000
ctcctgacct cgtaatccgc ccgcctggc ctcccaaagt gctgggatta cacgcgtaag 18060
ccatggcgcc cagcccatgt ggccatttt cagttagaga agccagaggc ccatcactct 18120
cgggtgctcc ctgggccatg ctctgcctca gccagaagca ctgagggaag gtcagcctcg 18180
gcccttgccc cagccacagt cacagataaa ggggcctgca caggtctgtg tggctccaga 18240
gctcgtcacc caacacacga cgcttccatg tgaatagccc caggtgcatc atgaagagcg 18300
atggccgctg cagaggcaga agaatccgc ggggaagcag gtgggagaga ggctgagaac 18360
agaccagacc ctggagctac agaccctatg ttccaaccct ggctgggact agctgtgtgg 18420
ctctgggcaa attcatatgc ttctgtgctc acaggggatc aaaatagcaa acacaggcta 18480
ggcacagtgg ttcacaccta taatcccagt gcttgagag gccgaggtgg acacatggct 18540
taagctcagg agtttgagac cagcctgggc aacatgggtga aacctcgtct ctacaaaaa 18600
aataccaaat aaattagcca ggcgtgggtg tacgtgcctg tggctcagc tacttggaag 18660
gctgaggcgg gaggaacact tgagcccaag aagtcaaggc tgtggccgcg tgtggtggct 18720
cacgcctgta atcccagcac tttagaggc tcaggtgggt ggatcacttg tgatcaggag 18780

ttcaagacca gcctggccaa catggtgaaa ccccgctcct actaaaaaaaa tacaacaatt 18840
tgccaggcgt ggtggcgggc acctgtaatc ccagctactt gggaggctga ggcaggagaa 18900
tagtiagaac ttgggaggtg gaggtttag tagccaaga tgggtgccgt gcactccage 18960
cagggggaca gagcaagact ccatcccaa aaaaaaaaaa acaacaac aaacaaaaa 19020
agaggtcaag gctgcagtga accatgattg tgccaatgca ctccagcctg ggtgacaaag 19080
tgagaccctg cctcaaaaca ataaaaatat aaataaaat aaaacataat agcaaactt 19140
tcatagaggt ggtatgagca ttaaatgaac tgataaacgt ccctggaaaa cagtaagtgc 19200
tatggaagga ttgctgccg ccaccgccac caccattagc atgtttcaac ctccatcacc 19260
ctcactgtcc cctgtacca tctttgacc agggcactcc cagctgcagc cttctatcc 19320
tctgtccac cctcataac tgtaagatca ctgagctccc aagaaccaca gtctacagg 19380
taaccacatt tccaaatctc aaaccagacc cgctggtctg cacttcagg gacaacagga 19440
tattttcaa ccagcccaa agagatgtgt ggctcagcat aagaggaaca ggagaaactg 19500
aggcctcttg ccctgagaat gagcttgga gtggatgtcc cggcctcact caaacctca 19560
gatgactgag gccagccag gagcttgagt gtaccctcag gtcataccct gagccagaag 19620
caccagcta atccactcct catcactgac tccctccca taaaaaacct gtttgctgtt 19680
tcaggctgtt aagtgtggg ctgtttgtt acacagcaat ggataactaa cacacgaggc 19740
ctggcaagtg tggagcaaag ctgcccaagc ctcaagtct gttcatgtgg gtgtggcct 19800
gtgtttgcag aaatccagcc actgagtcct cccatgcagt cactactgcc ctctgcacag 19860
acacctgcca catccctgcc tgggccagga gctccactag tgcaggaatg gggctgccg 19920
tcccaggagg atccctgaca cctagcacag ggctagcagc aggcagcact tggtagtga 19980
ataaactgcc cttcacctgt acacagaagg gatgtttcta taaggggtaa ttaagtacag 20040
agctgggaag ctatgtgac cagaaggctc taaaagcaat taaccaacga ggggaaaacc 20100
cttcctactc attctggcc catttattg agcactgacc atgtggaagg cccctggtg 20160

agactgggga atgcaccaat aactgagaca gcttccggct gtgacctca ggatgcctga 20220
gctgggtag ggccagggtg ggggtggtgc gtgtgacagg gttactgtc acaacctgc 20280
cgggccataa gccctccca acaattcaa aatccaaaac gctctgaaga tggaaagctt 20340
ttgtgtca tctggtgaca aaacctcatt tggatcatgg gccgggtgcg gtggctcacg 20400
cctgtaatcc cagcactctg ggagccgagg ggaaggatcc ctgagctta ggagttag 20460
accagcctga gcaacatgtg agaccccgct tctacaaaa atacaaaat tagccagggtg 20520
tgggtggcgca ctctgtagt ccagctact cgggaggctg aggcgggagg atcgcttag 20580
cctgggaggt gggggctgca gtgagctgag attatgacat tgcactccag cctgggtgaa 20640
agagtgagac tctgtctca aaaaacaaag taaaaaaaa aaaaactgtg catgggtgtg 20700
ggctacagat agtctttct gccctactta gaatgaacgt gccacattg ctatagaaat 20760
attcaagggc tgggtggcaa tgccacacag accctgacgc tgtccaagt tctgagaagt 20820
cctgcattcc tcagggtccc agagttcag agaagagtct gtaggcctga gtaagaagg 20880
aacgccttca aaagccctgg ggacaaagg gaaaggggtg cccagagact gcgtgggtac 20940
ctaccggaac gagccgtcca ggtggcacg gtggatgaag ctgagcttg cgtcagccca 21000
gtagagcttc tgctctcca ggtcgatgt cagtccattg ggccagtaa tgcagagtc 21060
cacaatgatc ttccgggtgc tgccatccat ccctgccgc tcaatccggg gcgtctacc 21120
ccagtctgtc cagtacatgt acctgtgacg ggggcagggc aagagaagca gtaacacag 21180
atctgtttt tgtttgtc tgcatagat gagacatgaa acaacagaca gtgaactgc 21240
cctaaatct cacccatcg aaataacaa caggatgtt tttaggtatt cctgcctaa 21300
gctgggcaat caaatatac tattccaac ttttctcag ttaacagtaa attctgggca 21360
cctcccttc ttgtggtag aaagattcct tttctttg atgattgcct agtgactct 21420
gctgtaagtt tttaaagaa cttcaggta ttctgattt tttgctacc atgaaatgc 21480
tgtaaatgaa cctctaaaag gcaattcaaa acactcagga tggaatatta ttagtgga 21540

taaagaaatg agctatcggc tgggcccagt ggctcacacc tctaatacca gcactttggg 21600
aggccaaggc ggggtggatca cgaggctcggg agatcaagac catcctggct aacacagtga 21660
aaccctgtcc ctactaaaaa taaaaaacat tagccaggcg tggtagtgag cacctgtagt 21720
cccagctact taggaggctg aggcaggaga atcattgaa cccgggaggg ggagggtgca 21780
gtgagcagaa atgcaccat tgcactccat cctgggcgac agagcgagac tccatctcaa 21840
aaaaaaaaa aagaaaagaa aagaaatgat ctatcaagcc atgaaaagac atggaggaaa 21900
cttaaatgca tgttagtagg tgaaagagcc aatctgtatg agtcagttc taaacactct 21960
ggaaaaagca aatacacaga gacagtaaag catcagtggg tgccaggagt tggagaggag 22020
agggatgaat gaggggagca cagaaaatca gggcagtga actatcctgt atgacatgga 22080
atggtgggtg catgtcctta ctcatctgc taaaccaaga atgtacaaat caagggcgaa 22140
ccctcgtgta aacgtggatt ttgggtgatg gtgcgtcagc cagcttcat cagttgtaac 22200
aaatgtacca ccctgcacag gatgtgaca gttgggaagg ctgtgtgggt gtgaggacag 22260
ggatgtatag gaactcagta cctgctgctc atcaatttg ctgtgaacct acaactgtt 22320
gaaaaaatta agtctattta aaaacaacaa aacatggcca ggcacgatgg ctgcacctg 22380
taattccagt acttcgggag gctgaggtgg gtgggtcact tgagccaccc tgggcaacat 22440
ggcaaaatcc cacctctaca aaaaaataaa attaaaaaaa agttagctgg gcatggtggc 22500
acactctgt agtcccagct acttgggagg ctgacgtggg aggatccctt cagccctggg 22560
aggctcaggc tgcagtgagc tgtgactgta cactgcact ccagcctgga tgacagagt 22620
agaccctgcc taaaaaaaaa aaaaaaaagg ctgggtgcgg tggctcatgc ctgtaattcc 22680
agcgctttgg gaggccgaga tgggcggatc acgaggtcag gagatcgaga ccatcctggc 22740
taacacggtg aaaccccgtc tctactaaaa gtacaaaaaa aaaaattagc cgggcatggt 22800
ggcggacacc ttagtcaca gctactcggg aggctgaggc aggagaatgg cgtgaacccg 22860
ggaggcggag ctgacagtga gccaaatca caccactgca ctctcagcct gggagacagc 22920

aacactccgt ctcaaaaaaa aaagaataaa acccatggct gggatggacc ctgaacctgc 22980
agctgcagct gticctgggt aggtctgtgg gcgacgtggc ttgcttctc catgttccca 23040
agagacaagc atcacccatc catgagaaac aagcacatcc tcagggcgcc cttacgtgat 23100
ctctggccaa tgaaccaaga caaagtgagc agacaccagg tctgggatgg caggtcccac 23160
ccccaccagt gccagtgtg cctgttttg aggtgaccac aggggtgtgtg cccagaggct 23220
gggcgtgact ctacgcggag accagagggg aaccacacca gcttgagga ctcaattccc 23280
atcccagcca gctgggatga gccacaggac acaagggtg gcagacctat tgtgtttgt 23340
ccacccttca cagcagagaa aggggacagt gccagaatg tcctctgagg agcctcctcc 23400
cactcttgt ccttgtaaaa tgggtgtgac tccctgtct ccttcttct ggggtggcg 23460
gcaaacccca tccccctcag ccttagcaag tgatttagaa acaggcagct cgcccaagcc 23520
aggcatgaga gtgatcccg gacacaggga gaacaagccc cgctttgccc tctgggggtc 23580
tccattcagc agaagaggca aatgacagac acacagccgc ctctcccc accatgggtc 23640
tctgcagcct caggagcctc aggtgcacca agggccacc catccagggg gccatgcttc 23700
cttgagtgt atcgttcctg agcgagtacc atctccacct tccagagggg ctgtgacaag 23760
atcaacaaga atgagggcac aggagcctcg aaccaaacat gccctcttc ctgcagaggc 23820
tgactgcgc cagctgctat caccaagccc ctgctctcc ggccccgtgg ggacagggtg 23880
agaggggtgt cacatggaac agctctcaa acagtccctc tcaagctgct gtctcctgtg 23940
catctagtga gaaccaacc acaaaggga aggtgggaat tgctattccc attaggcaga 24000
tgagaaaact gaggccccga aaggctggcc tgttcaggt tacaggcgct gagcggctgc 24060
tctgggaaca cacttggtgt ctgctgagg cccgagcccg gccatcatat gactcacct 24120
tcgccagcaa agcccgggtg tgggtgaact ttcttgga gcctgggact ccaaggtgct 24180
ggcagccagc ccagggaagg ctcccgcgtg cctgcggcag acgccttgct ttacctgcac 24240
gtccccacc ctaggagcct ggacagagcc cagaccctcc gccacctct gagaaggtat 24300

caggggcatc agtctggact tgggggggaa tccacacagg ccttcccaa atgctccacc 24360
gtggcccatg gaaaaggctg gaaaacgtgc aggagcagga gcctccgcat ggagcataat 24420
tcacattcct tccccgagtt tcataacaga ggcctgctgg tttcctaaa tggggaattt 24480
gcgagccagt cggtgaccag agactggttg gcgtggacgt gctcttcag agtctcaaac 24540
gctaccacaa gccagccaa attccacgga ggaaaatcga ctccgaaga aaagagctgc 24600
agcatggcct tcgtgcagag ccagctgcgg ttgtggtgt gtgtatttt agggaagggc 24660
cattttgcat ttaaagagg gggttgggtt tcacctggc ttaattga gaccggggg 24720
ccactgcagc ccctgtcag gctggtacag gccggggact cctccatgc taagccagt 24780
tctttctggc ccagatcct caggggccag aggtcatcc ccagagccc ctctgccacc 24840
cacatgggta ccctgggcct gggagggatg tgccttcct caacctgcc tggatgtccg 24900
cacggggcca cctgcattgc tgaaactga acgaagtcga gtctcaggag gggccccct 24960
ggctgcaggg ctctgatcc ttgtggcac gtgcacactg aggtggacgc tcggaccag 25020
agacccctt catgatgat gccggggcag gaacccctc ctctaggaa ggacctgt 25080
gggggacagc actgcaggag ggcacaggag atgacgggg ctctagcagg gccgggagga 25140
aggccaagat gtcctcgca accgtgtcc tgtggccagg acagaggaca aaccaccct 25200
ccactgtccc cactctcagg acagcagtc tggcccagga ctacgcgcc acattatgc 25260
ctgaggacca ctattcaagt cagtatttg cgagcagggg ttgctgccgc gggcgctgtg 25320
acaggctgga atcctctccc tctcctctc cctctccgga gacatggagc ctacaggac 25380
agagtcagca cctcagggta ggacctggc tggcgtcatc agcatcactg gatctgatga 25440
gtgggagccg gcatctcact gtttcactc tctattcaa atgactggag caaagggaag 25500
gtgtggggag aggcccagga atcaacta aggtcaact tggcccagg ggcagggtg 25560
ggagtgaaca gccacaggtg tgatctggg gagggcttct gggagagaat tcagaggca 25620
gcatgtagag gaaccattc aaatagttaa gaaaagccag agccaaacag ggacagtgg 25680

ctcgcagaga tgatgcaggc aaagccagct cagatctgag catgggaaag actactcca 25740
accaagggcc cagcatctcc caaccaagca ccaagtacct cccaacaaa tgccaagcac 25800
ctccaatca aafacctccc aaccaagcac ctgacacctc tcaactggac accaactact 25860
cccaaccagg caccaagtac ctccaacca agtgccaagc acctcccaac caagtaccaa 25920
ttacctcca accaagcgcc tagcacctcc caactgagca tcatgcacct cccaacagag 25980
catctagcac ctcccaactg atcacctccc aacctagcac cgagcacctc ccaaccaagt 26040
gcagagcacc tccaaccaa gtgccaagca cctcccaatc aaatacctcc caaccaagca 26100
cctagcacct ctcaactgga caccaacaac tccaaccaa ggcgaagca cctcctaaca 26160
aagtaccaat cacttccaa ccgagcacct agcacctccc aactgagcat catgcacctc 26220
ccaacaaatc acctagcacc tccgactga tcacctcca acctagcact gagcacttcc 26280
caaccaacat agcaaaagcc ataaagaagt aaaaagacaa aaccacgtag gcatggagac 26340
tggaactctg gtggcgagga aagggcattt ttattataac gacagctaac atttgtgaa 26400
ctcacaact gtcttggtg tttcctcat gacatgcagc atgtcacgc ctctgtacag 26460
acaaggatac tgaggcacag agtggcaccg tgccaacctt gtctcatctt ttatcgaac 26520
ctacatgcag agtgccagca aatccagctg tctttctct tcagaacaga tcccaaatct 26580
cgccactcct tcccccaaca agtgagggtg ccccgctgct gctttctgtc gccaggatcc 26640
cggtataaac cgtggagagg gctcctgccc ccacgccacc caccacacag ctactctcg 26700
ctccagccac caggggatgc ctccagcac gagtcaagc tggcacctcc tctgtcgag 26760
acctcatgtg tctctctc acacctggg cctgtttcc ctacattctg ctacagcccc 26820
tcaaacaggc cccgccccaa accagcccag ggccttgca ctggctgac cctctgctg 26880
gaccgcgtg cccccagaca gccacacggt tctcagctc atctgcttcc agtctcgact 26940
caaaagtcac caagaggcct tccagcacc tgagctccga cggaagcccc tcgccacagc 27000
accaagcac tgctttatcc ccctacgcac acgtccctt caaatactat tcattacca 27060

tctctccca ctactgaaa gggccagaga ctgggctata cccgctgctg ggggagcagg 27120
accaggcgca agggctcaca aatgcagtgg atgcctgggt gggaggtgag ggagctgcag 27180
cgaccacgc tgggagggaa cgcaatgaca ggaggagcgc aggtcctggc gacacgatgg 27240
ccatggcagc cgctggtgag caaccgcagg ccggccctgg gagagggctt ctagcaagct 27300
gctatcttca gcctctccga ctactgcaga tgccccctcc tagccagaga cactgctaca 27360
ccagccgacc cticcaaaaa gaaggtcagt aacccgcga ctctggagc cacagtgcag 27420
ggggagaggg ctgagagggc aacagttcac caagcggaac agaggctgcc ccggaggtca 27480
gctggctccc cggcagctgc aggggtggct agcccactcg gagggcagcg agggcatacg 27540
aggggctcca gggatgagt gttgccagc acagcacccc tgggaggccg ggggcacttc 27600
tcaggtagtg ggggcacgag gctgctctgg cctgacctca gggactcaa atactttggc 27660
gataaattcc accgtgtccc acccctgctg gtaccccata ctacacaca gactggttca 27720
gatgcagaca ctctcgca catactcgt cacacgggca catacagtg cacacacagt 27780
cacatgcgca cactcataca cacacaaata tccactcaca cgcattgatg cacacacag 27840
gacacacaca ggctcacag tatgcagca tatgctgca cagcacaca cacacacaca 27900
cgctcacatc ctccactcc cacactcagt tgctcagaca cacacagcc tggctctcac 27960
acaaacctgt tgggctctga aaggctccag ccttcccat gctcgtcaga agccagtcaa 28020
tggcttccta agtcaccaca cagatcaaag aggtgaactt ggccacatgg cactctgctt 28080
cctgagctcc caaacaccag ccttggtgag gacagacct caccacac cctattccc 28140
actaccctgg gcaggccag aggggggga tctgcaggat ctggcaacca gccctccc 28200
cccggctcct gcagccggca ccatgggagt cagggggagg tctactgaaa gggcaacagc 28260
aagtgggtgg cccaggact agagccagg ggtcttcagt cctactccag agcttgaca 28320
ctgtcccaca gggcatggcc aagggaagg cttccagagc cctgacttca gggaggagg 28380
caggcgggct cctgtggcag gcctggatgc atggccgcc actcctggga ctttctaacc 28440

tagaatatct aggtcaggct ggggtgcagt gctcacgcct gcaatcccaa cactttggga 28500
ggccgaggag ggtggatcac ttgaggtag gagttgaga ccagcctggc caacatggcg 28560
aaaccctgtg tctactaaaa atacaaaacc tagccagggtg tggtagtgca cgcctgtaat 28620
cacagctact caggaggctg aggcaggaga atcacttgaa ctcgggaggt ggaggttgca 28680
gtgagctgag atcgtgcat tgcgcaaaga agatctaggc cggcccctca accggtgagg 28740
tccaggctgg gagtgtgag agactgtgtg gacactgaat gaactaacag gcaaagggt 28800
tccaactgag cctgggggtg gtgggaaatg gctcttgtgt tctagtcaag acctctgcca 28860
accagttctg aactgaccc agcacagaac ctgacaggtc agcaagggcc agggcttagc 28920
acagcccagg taagggtgtg tgtacggccc ccagagtcac tcccaggctg caagaaaagg 28980
gacaaaggag ggacaagggg tggccaagca aactgttccc tctgctcggg agtctgggt 29040
gacctggcct agctggccag tggagctggg ccacctcccc taaactctc caccgagac 29100
ttcgactcca aagctttcct gccaccacg ctctccccac ctgggatcac ggccaggccc 29160
tgagccttca agggcccagg tgaactcagc cagactagga gctgaggagg acacagggca 29220
gcttcagaa cggacccgag aaccactccc agcaggttct gcttcagac aaggagctgc 29280
acttttcag ccaatgcaat tagaaagcca ggagaagggt caaattccac ctgcctgagc 29340
gtccgcactt ccaggccgc ccaccataca cacagcaaag atgtgtttaa ccattcaaac 29400
ccatggccaa ccacatcgtt tgcctcagac atgcaagttt taaaaggaa cataactatg 29460
ggccaggcac ggtggttcac gtctgtaac ccagcactt gggaggccga ggtgggtgga 29520
tcacctgagg tcaggagtic gagaccagcc tagacacat ggtgaaaccc catctgtacc 29580
aaaactacaa aaattagctg ggcgtggtg tgggcgcctg taatcccagc tacttgggaa 29640
gtgaggcag gagaatcact tgaacccggg aggcgaaggt tgcagtgagc cgagattgtg 29700
ccactgcact ccagcctggg caacaaggga gactccatct caattaaaaa aaaaaaaaaa 29760
aaaaaggaaac ataactatgg agtctcaagg ggaagtaatt ccttcaacaa taacaaatct 29820

tgaaagctga gctctttttt ttttttgaga caggatctcc tcacttgtc gccaggctg 29880
gagtgcagtg gtgggatcac agtcactgc agcctcgatc tccaggctc aaatgatcct 29940
cctacctcag cctcccaaga agctgggatt acagggtcat accatcacac ccgattcatt 30000
ttgtatact ttgaagagat ggggtctcac catgttgccc agtgtggtct tgaattcctg 30060
gactcaggtg atctgcccgc ctggcctcc cagagtgtg ggattacagg cctgagccaa 30120
cacccccacg ggttcatttt cagagtcgca ccgagtgtg gggttacagg cctgagccaa 30180
cccccccacg ggttcatttt aagagtgaca ccgagtgtg gggttacagg cctgaaccaa 30240
cccccccacg agttcatttt cagagtcgca ccgagtgtg gggttacagg cctgagccaa 30300
cccccccacg ggttcatttt aagagtgaca ccgagtgtg gggttacagg cctgagccaa 30360
cacccccacg ggttcatttt cagagtcaca ccgagtgtg gggttacagg cctgagccaa 30420
cccccccacg ggttcatttt cagagtcaca cctttttct gaaaaacaac ttgggctcat 30480
gcaaattcga gagagagatg gtgacactcc ccgccccctg gaccaggtg gagtgcgagc 30540
agggtttacc cgtgagcggg gtccaaggcg atggccctcg gctggtcaag gtcctgccag 30600
aagagcacct tccgggatgt gccattgagg ttggccacct cgatgcgggtt ggtctctgag 30660
tccgtccagt acagttctt gccacccag tcgcaggcga ggccgtcggg agagaccagg 30720
ccggagatga ccacgttctg cacggcggcc cccgtctggt tcaggtaggt ctgcttgatg 30780
gcctctcgc tcacgtctgt ccagtacag gctcccttg aaaactggaa gtccactgcg 30840
gccgcacct ccaggccgct gaccacgatg gtggactcca gcttgactcc gccggcgtcc 30900
accagccgta cgtcccggcg gttggcaaat agcaggagcg gcgaggctgt ggggcagaag 30960
caaaccgtga gggccactgg ctaagccagc aagatacaca gccctgggat ggagcactat 31020
gccagagca ctctggtac tgccctgccc atgccaaga cctccagttc ctctctcca 31080
cccctaaggc gttgtcagga agttgcctgg gcagccccgg cccgcatcat tcagaggctc 31140
ctgcagcgca gcaaacagcc ttctccac attcgtgac agcacctgtt tgtttacaa 31200

ctgttacgtc tgttccccca gatatgggtg acccttcctg ccatgcccaa aacctccac 31260
atcgtcctcc agaggctaca ggggccctgt cctgttctgc agagaagcca catcccctt 31320
gttggcctga cacaggggat ggggacatgc aggcacagca ctggccatgc tgctcgtac 31380
agaccagcc acagggccac atttttgag ggggtcagag cccaggccag acagagcctc 31440
aagattcctt tacaagtctt tgaccactgt ccaagctcag gcccgtttc ttggccgtgg 31500
catcagctc ccatccacc ctgtattcca tgtttctcc accctgctc tggacattcc 31560
tacatttaa gggtcactct ggaatgccac ccttggctc agacacctc cacagctccc 31620
tgtgccagt ccatgcagaa caaggtcaga cccctagcc tggcctcaa ggccttgcc 31680
tctggcctca cctacactc tctccaccac cccacccaa gcattcctga tctgcctgag 31740
gccaggctgg ctccctcacc tcctgtgca cgcagccct cagcccctc tgctgtgca 31800
agaagcctca tctcacagac aacggtctca tcccacaac gggctcaatg agaatcagg 31860
agaggcctc agaccatcac cccaccagac acctcagacg tcggaccagg agggccagc 31920
aaccaccaac acagactcag agggactaag aagccacatg aggagtgaac acaagatgtg 31980
gacaggagga ggtaagggc ctccaggag ctccatcagt ccgtgtctg ctgtcagcag 32040
ggtaggctg ggctggccac aaacacccc aaaaaacatc tgaagcctg gcttgaaca 32100
gctgacattc ctcatgaaa ctgcagacc ctgggtcctc ctgcgcagat gggggagccc 32160
agccaacccc aactccac cttaccaag aaagagaaag ccaaaacaaa ctcaactcag 32220
ccaatgacaa tcacagaact gaatcctgta gttagttcag ttggttcat ttcagcaggg 32280
gaaagattg cagcctctat gaggttagct gggaacacaa agggccagag catggcccag 32340
gagaccccag cgcagtgggg tagatggctc cgagcacagg cctccctgcc aagacaagca 32400
ctggctcaaa tcctggcccc tccattccc aggagacatg ctccacagga tgggaggaca 32460
cacagaggac ctgaggccag gaaaatgaca gcggcgctc cgccgcccc cccgtgctgt 32520
catcatctta ggtctacagt tctttgtgc aacgaggac actgtgaaag tcaaacaaca 32580

ggaaggcata ggccacaaat aaagacaaac gggacttcat gggaagctaa agattttgtg 32640
catcaaaaga cactatcgag agagtaaaaa ggcaaccac agaagagag aaaatattc 32700
caaatcatag atctactaag agattaatat ccatgaaata cagagaactc ctaaaactca 32760
acaatgagaa aacaactaag ccaactcaaa aatgggcaaa caactgaac agacatttct 32820
ccaaagatga catataaatg gccataaac acatcaaaac aggcttaata tatccctaata 32880
catcaggga atgcaaatca aaactacaat aagataccat ctgaccaa ttaggacggc 32940
tactatcaaa aaaacaaaat agcaagtgtt ggtgaggatc tggagcaact ggaaccctg 33000
tgcaccactg gcaaaaatgt gaaatggtgc agctactatg gaaaacagca tggcagttcc 33060
ccaaaaactt aaacacagaa ttaccatatg acccagcaat ttcgcttgg gttatatacc 33120
caaaagaact gaaaacaggg acacaatcag atatgcatac acctggatc acagcagcat 33180
ccttccaac agctaaaaca tggaggcagc caggcatggt ggctcacgcc tgtaatcca 33240
gcactttggg aggctgaggc ggggtgatca cctgaggatc ggagttcgag accagcctgg 33300
ccaacatggt gaaacccgt ctctactaaa atacaaaaat tagctggcg tagtgacggg 33360
cacctgtaat ccagctact cacaagtctg aggcaggaga atcactgaa ccctggaagt 33420
ggacgttgca gtgagccaag attgcgccac tgcattccag cctgggtgac acagcgagac 33480
tctgtctcaa aaaacagcaa acaaaaaa aaaaaa caaacatgga agcaaccaa 33540
gcgtccctct actgaggat gaatagcggg gcaaaatctg ctccatccac acaatggagt 33600
actattcagt ctcaaaaagg aaaaagattc tggtcaggca cggtggctca tgcctgtaat 33660
cccagcactt ggggaggctg aggcgggtgg atcacctgaa gtcaggaatt caaggccgc 33720
ctggccaaga ctggcaccna gctacacana aagtatang ccccgaaa 33769

<210> 9

<211> 72049

<212> DNA

<213> Homo sapiens

<220>

<221> unsure

<222> (8356),(8385),(38585)

<223> Identity of nucleotide sequences at the above locations are unknown.

<400> 9

tataccttgc gcggaccttc ggctcctgtg gtgaagacaa tatgaagaaa atagaaatta 60
cccataattt tgccacacag acttagttgt gtccatgtat cttgtgcacc tttttctgt 120
ttacggatca aaatcgactt ttagggtcag gcgcggtggc tcacacctgt aatcccaaca 180
ctttgggagg ctggagttgg ggttggggggg tggatcactg aagatcagga gtttgagacc 240
agcctggcca acatggcgaa actccatctc tactaaaaat aaaagattag ccaggcgtgg 300
tggtgggtgc ctctaattccc agctactccg gaggctgagg caggagaatc gctgaaccc 360
aggagacaga ggttgcagtg agccaggatc acgccactgc actccagcct ggcaacagag 420
cgagactctg tctcaaaaaa aaaaataaaa ataaaataaa taaatacata aattgacttt 480
taggagattg gticaaacia tgtgtgtaat gttgtgtctg agtgttttc atttatcgtt 540
catgcaaatt ccgacatcat tcactcttct ccagagtgtg ctgttttct gcctgtgtca 600
tcacccgtca cctgaatgc cctcgtttag gtaaaataag tacattttat tcaaaaatat 660
ttgaggacat ttgggttgc tccaggttct tggcttgag tttgtctgtt cttgtggagc 720
catggtggtg tctggttga ggaacctcca tgcgttcag ctgctgcttc tgcctgtgtt 780
cttagagagg aaatgctggg gtccgcggtt cccgggctgc tgaccaggaa gcctgcggtg 840

ctttacggcc cttccagaag cgggagatgc cccacttaa gtgtcagaca ggcctttcca 900
cctcactggc agctctgagc ggctcccttc tatttgcaga tgactgagaa gttaccaatt 960
tccacgttta ctgactgctg ttctctctgt taatttgtat ttatagtctt cgctaattta 1020
ttgctagggt ttggtgttg tccctattga ctgtatgcc tttaatttt taaacaaca 1080
ttaatatact tcatttttt agagcagttt taagtttaca ggaaaattaa gggacaagta 1140
cagagagttc ctccacctg ctgtcctct ctctcctcc ccaccttccc tcttccct 1200
attgtaactt tctttctgat attataaaag tcactcatgg ctgggcgtgg tggctcacgc 1260
ctgtaatccc agcacgttgg gaggcagagg caggcagatc acctgaggtc aggagttcca 1320
gaccagcctg gccaacatgg tgaacccccg tcttactaa aaacacaaaa agttagccag 1380
gcgtgggtggc gggcacctgt aatcccagct actcaggagg ctgaggcagg agaatggcgt 1440
gaacctggga ggcagaggtt acagtgagtc gagatcgcg cactgcactc cagcctgggc 1500
aataagagtg aagcttcgtc taaaaaaca agtcacacac gcttcttgta cgagggtcat 1560
ttggccgagg ggccagatgg ctccatctt agttgggaca ggccatgagc tcggaatgct 1620
ttttacatat ttacatggtt gagaagaaaa tcaggagaat aatgttttg gacatgggaa 1680
aatgacatgg aatttgcatt ttagtgtcca taaatgaagt ttgtttgct cccagctgtg 1740
ttgactgagg caggctggct tctacagct gcggcagagc tgaggaggcg ggaaggagac 1800
cgtgcaggcc gcagaccga aaatatgtc tctctggccc tcccagagt gcttgccgac 1860
ctctgtccga cagctagaag gaaggatagg acccgtccga cgataaccac tgttgacatt 1920
tgagcgcgtt tcttcccgg cttttgtgtg agagtggcag tctgtttgct ttgtggtcg 1980
ggatctgctg cacgcacggc gggctgtttg catgaggctt cctggaggat agggctgggc 2040
tcggagctgc acgcagtggg gcgtgtctg catgcagtgg ggcctcagaa gagagctgtg 2100
gtgggcgggg cagtgccaac gctggtgggt gccaggcctc cacgctcaga tcagccccgg 2160
cgacaggttt gggccaccct ctctctggcc tctgtgcagt ggcccaggcc gtctgtctg 2220

cctggcacac ttgcctctgt cctccactg aagcgctcct cttaccctct gctcccggct 2280
gggtacgttg aatttgttcc ctcaaggaga tatgctaaag gtctaacccc aggaacctgt 2340
gtatgtgatc taatttgaa acagggtctt ggctgatgta atcaagcgag gatgaggtca 2400
ccctagagta gggggcctat atccacgggtg ctgggtgcct catgagagca ggtgagcaga 2460
cactgacact caggggtgaa ggctgcatgg agtcagaaca gggttagtg cgatggcggc 2520
cacaagccaa ggaactccaa gtatttctg caacaccaga agctggaaga tgccaggaag 2580
gatcctgccc tggagccttc ggagggagtc tgtccctgca gacgtctga cttttgattg 2640
cagggatgca tgtcttaggg tgtgtggggg ggtgcatttc tgatgttaga agccacctgg 2700
ttggtggcga tgtgtcacgg gagccctctg caggttctgc gtgtccatgt ggtcggggac 2760
agagggtggc agggacggac ggtgtcgagc tggacatgc catgacgtcg gccatccctt 2820
gggatggctt tttgtttt aggataaggc tgcctgccag gaagctgtgc cctgcctggc 2880
ccttccccca agcccctggc ctgtgcttgg cctcgcggaa gggatgtcgc ccttctctcc 2940
tgcatgcgtg cagggaggaa ggggagaggt cagcagcccg cctggaggag gctcgggcga 3000
ggggaagggt tcacttcag gcaatgtgt ggggctgtt aaacaacccc aaagaaaacc 3060
atttgccaa actgttagt tccaaacatt ttacttctt ggtgtttaaa taaattccta 3120
ccaagactct gtagctggc ccaggaagg agttggcctc tcttcttat agcccggcac 3180
agtcagtccc ctgcacctgc cctcccaac ccaggcctg cttcccgtg gccatggctg 3240
ctgcccggac ctctctacac acagaacctc ctggaggcca gctgtgggca ccagccttgg 3300
cagggctgtg gcggagccca ggctgctggt actctctctg cagctgtccc ctgctggcct 3360
ggctggacag cgtccccacc accactgggg tcacctctgt gctggtcaca gtcactcag 3420
accttcaggc aaatgggttg gatcctgcct ctctcccagg tgtctcagtc tctgaaaac 3480
tcaaaaacct cagaggcctt gcagcctgag ggggtgcaga gacacctct tcgaatcagt 3540
aaacacctac agattacccc cagcagtgaaggactgctt cgccacagag gtttgattta 3600

ctcctaagta attggaagg atgccgagaa taggttcctc atggtgggac tagaggccct 3660
ctgctgacct agttaacaga gggctagggc tgggtgtgct cagccccga aggttctagg 3720
cccatttggg acaccccgcc agaacctgcc acaacctgcc atgtggtgac agctaccta 3780
atcccagagg ctcttgagct ggagagcaga cctctcaatc tcagcaggcc cccacacag 3840
accccataac cctagtctgc cttcacagta cagttcgtgg ctatgtgttc acggatggtg 3900
ttgttcacct aaggtctctg ccctgtgacc ccaagggcgt cctgagggca gattccaagt 3960
ctgtttcgtc caccctcct tcctagcag cgggtccagg gcctggcctg aactagcttc 4020
ccacagagat actggtggga tgatgaaggc agccaggcgg caagtgaana acgcacttc 4080
tgcattgtct ggctcctggg attgaagtgt ttgaggaagc aaagtgaagt gagcttctct 4140
cttgcggctg tgtgtccttg ggccgggagc ctaccctctc tgagcgttgg ggtccttgc 4200
agtagaatgg ggcctcctca tagctcaagg ggtggtgtgt gaaaattgtg ctattgtgtt 4260
actttaatga tttttttt ttcagacaa agtctcacc caacgcgcag gctggagtgc 4320
agtggcgga tctcagctca ttgcaacctc tgcctcctgg gttaagtga ttctcctgcc 4380
tcagcctccc aagtagctgg aattacagga gtgcgccacc agggccggca ttttttcta 4440
tttttagtag agaggggggt ttaccatgtt ggctaggctg gtcttgaact cctgacctca 4500
ggtgatccac ctgcctcggc ctcccaaagt gctgggatta caagcatgag ccaccgcgc 4560
cggcctactt tagtgatttc ttaggaggac agaggggaac ggctggcaag acaggcttgg 4620
aatgtgtttt gggatcaagt gccggttct gtctggcact ggcgttctct gtggggccat 4680
gatggacaca ctgctgaggt caagcgtgat tcgtcttgcg ctgtgcctgg cagtctcatt 4740
ggaaagtctt gtagacatcg tgtgatggg gctcttccc gccaaagcct tggggacctt 4800
ccaggactgt gatcctcca cagtggctgt taagcaggga ccttcgtga agtggagtct 4860
ctggtcccct ccaagtcata gctagacagg gactcgggca tcgccaagcc tggctgatta 4920
ttcactggat gaggagacag gccagagag gggcaggaac ctgcccagg tcaccagca 4980

ggccccagag gtttcgggtct cggattctcc ctgctcatcc ctggatgtag tgctgtgtg 5040
gatgtggttc tgtgtgggg gctgtggaga gcagggggct tgtgccagga cccagtgag 5100
ggtggcgccc tcgcatgag gccgactgtt ggtatggggc ggccatccac tggggtgtgg 5160
ggaggaacag ctttctgag gaggaggtgg cgggaggaac agcttcctg aggaggaggt 5220
ggcggtgctg tgtgacctgg gcctgaagg acaggtccat tgtcaacaga acatttggg 5280
agtggagcct agaggagaa aatttgtga aattcagatt cccctcccc taccaataca 5340
caccaaatca gatgcccctg accagatcta aattggctc tcagagattt ccattgtagc 5400
tgggcacttg gggaaccttc taagtgtgc ctctgcctct cccagcctg cctgcctcag 5460
tttccccagc cctgggcccc tgctgtgtt gccatcacgt gggcgccctc tagtgaggga 5520
atcagattat gcactccggg gcttgagca ggagtcagga ggggtcctg tcttcttg 5580
aaacgttga tgccgggatc ctggaacagt ctctgattc ctctggcg gaaccagagc 5640
ctgggcacag gggacctct gttgttgaa ggctgcagcc tggcaggga ctcaggagat 5700
ctggcagttg gctgcagggc caggtctagg ggccaggga tcaggaggc tctgggctgg 5760
ttcagccccg ggccccttg cagattgtga cctgggcccc tgtcagggg catggccaca 5820
ggatgctggg aggggtctct gaccctgacc ttctggctc tgtcatcct tgagaccaga 5880
aaggcttga acaaatgagt agacgatgcc ctaacctggg gaggagcca catcctgatc 5940
ccagcaacct cgggaaggat ctgtcaggat tatggggcac cctgggggcc ccaagtctgc 6000
atgggtctcc acttgcaatt tcttaggaa gctctgataa atccaaactg ggggtcctag 6060
gacacagica gaaatgctga taccgttggtg tgtggagcct cgggccctgg gggtcaggag 6120
catgtggagg gtgggccacg ggggttcaga agagaatcct gtaaccccc acccccaaa 6180
ctgaagccca cttgagggcc atggctgaaa ggttgggggg tctcgtgcg tctgtggag 6240
tgggtggtga ggagtccttg gtttgcacg cctctgggcc tgagcgcgcg gacccgtcc 6300
acagcggatc cctgggccct gttgctcaga tgctctcaga gtgttgctgt ggccacggag 6360

ggagcctgag ttaagcttct cttgtgccgg ttgtacgctg tcaggtcaca ctggtgagtt 6420
aggcaggggca cagatgcccc gagcagaggg aactttcctt ggggattcaa cacgtgcaag 6480
tcttaggggc tggcaaatcc tgccctcagc tagagagggg gcttttattt gagaccagaa 6540
tcacctgagc atcctcctgt cccagctgt gtccagcctg tctgcaggga catcctgaga 6600
ggaccaggct ctccctcat ccacctgcct aagtgcact ctgaaccctg tccacctgtg 6660
ccgtggaggg gcgtgacctc aagctgctca gccagcagca ggcttgccc tggggggcag 6720
cagagaccca ggtggctgtg ggggtgggtgc ttcgtggcgt ggttctgaaa cticgttga 6780
agtgtgtga cagtgccttg cctgttctct gtgggacctt attagaaac gaggtctgag 6840
ttactggggg tcatcactgt gttctgatgg ccagctgtg tggaggccgc ggtgcagccc 6900
catccaagga gccagggccc tgggtctagc cgtgaccaga atgcatgccc cggaggtgtt 6960
tctcatctcg cacctgtgtt gcctgggtgt tcaagtggc gtgaaactct gtgttagctc 7020
ttggtgttcc tgaaagtgc cccgggtctc aggcctcaga accagggtt ccttcactc 7080
cgggtggcctg ggagcatctg ggcagttgag caaagagggc gattcactg aaggatgtgt 7140
ctggccctgc ctaggagccc cccggcacgg tgctggggcc tgaagctgcc ctcgggtggt 7200
ggagaggagg gagcgtgaa gtggcgtcga gctgggcagg aagggtgagc ccctgcaagg 7260
tgggcatgct ggggacgctg agcagcatgg ccagcagctg ggtctgcagc ctggtacccg 7320
gcgggacttg tggttggggc tggtttgtgg ccaggagagg ggctggcagg agacaagggg 7380
gactgtgagg cagctccac ccagcagctg aagcccaatg gcctggctgt gtggctctca 7440
gctgcgtgca taaccttca gtgcttcagt tctctcattt gtaaatgag gaaacaaaca 7500
gtgccagcct cccagagggtg tcatgaggat gaacgagtga ccatgtagca tgggctgggt 7560
gcgtgtcacc taacatcacc agcctttgca aggagagccc tgggggcctg gctgagtatt 7620
tcccttgccc ggcccacccc aggcctagac ttgtgctgc tgcaggcct tgaacctga 7680
ccccattgca cctgtctcca caggagccga ggaggtgctg ctgctggccc ggcggacgga 7740

cctacggagg atctcgctgg acacgccgga cttcaccgac atcgtgctgc aggtggacga 7800
catccggcac gccattgcca tcgactacga cccgctagag ggctatgtct actggacaga 7860
tgacgagggtg cgggccatcc gcagggcgta cctggacggg tctggggcgc agacgctggt 7920
caacaccgag atcaacgacc ccgatggcat cgcggtcgac tgggtggccc gaaacctcta 7980
ctggaccgac acgggcacgg accgcatcga ggtgacgcgc ctcaacggca cctcccgcaa 8040
gatcctggtg tcggaggacc tggacgagcc ccgagccatc gactgcacc ccgtgatggg 8100
gtaagacggg cgggggctgg ggcctggagc cagggccagg ccaagcacag gcgagaggga 8160
gattgacctg gacctgtcat tctgggacac tgtcttgcac cagaacccgg aggagggtt 8220
gttaaacac cggcagctgg gccccaccc cagagcgggtg attcaggagc tccagggcgg 8280
ggctgaagac ttgggtttct aacaagcacc ccagtgggcc ggtgctgctg ctgggtccat 8340
gcgtagaaag ccctgnaaac tggaggggagc ctttgtccc cctgncttca gtttctcat 8400
ctgtagaatg gaacggtcca tctgggtgat ttccaggatg acagtagtga cagtaagggc 8460
agcctctgtg aactgacca cagtacaggc caggcctctt ttttcttt tttttgag 8520
atggagtctc actctgtcgc ccaggctgga gtgcagtggg gtgatctcag ctactacaa 8580
cctctgcctc ctgggctcaa gtgattctcc tgcctcagcc tctgagtag ctgggattac 8640
agggtgcctgc cactgtgctt ggctaattgt tgtattttg gtagagatgg ggttcaccg 8700
tcttgccag gctggtcgca aactcctgac ctacagtgat ccacctgcct cagcctcca 8760
aagtgtggg attacaggca tgagccacca cgcccggta ggccaggcct ctttgaaca 8820
ctttgcacac catgggtctt ttatccagg ggggtaggta cagttgtaca gttgaggaca 8880
ctgaagccca gagaggctca gggacttgcc cagggtcaca cagcaggatg tggcagggtg 8940
ggggctgggc ctggcagcgt ggctccagct ttccagcata gaaatctgtg aaagcagata 9000
gtttgtcggc cggtagggga gactttctga gaccgcccc agcggctcag agggtagtag 9060
ccaggggcct tctgggggc tcataacca gaacactgaa tgggaaaacc ctgatggagg 9120

aggcgtagtg gagctgtggg tgccgatggg aagtcacaga ggagctggga ggtcagtagc 9180
ggtgtgtccc tctgtggagc acttagtggg caccaggtgt gttccaggt tcatggccct 9240
gggacctgaa gtcagaagg tgaagtaact tgcccagggc accgctcggg cagcggcggg 9300
cagaggattt gtgggctgtg gagcctgtgc tcgtggccca gccctggggg ttgtgagtgt 9360
gctggccggg gagctttcc tgcaagtga ctggtgtcta ggagccagca tgcaggcag 9420
caggcagcgg gagtgcagca ggcagcggga gcacagcagg cagagggcgg ggctcagca 9480
gccatccgtg gacctgggg cacggaggca tgtgggagag ggctgctcca tggcagtggc 9540
tgaagggtg ggtgtgtccc cgaggagggt ggatgagggt aagaagtggg gtccccaggg 9600
gctttagcaa gaggaggccc aggaactggt tgccagctac agtgaaggga acacggccct 9660
gaggtcagga gcttggtaaa gtcactgtct acatgggcct cgggtgtctc atctgtgaaa 9720
aaggaaggga tggggaagct gactccaagg cccctcctag ccctggttc atgagtctga 9780
ggatcccagg gacatgggct tggcagtctg acctgtgagg tcgtggggtc caggaggagg 9840
caccgagctg gaagcgggag gcagaggggc tggccggctg ggtcagacac agctgaagca 9900
gaggctgtga ctggggcct cagaaccttc acccctgagc tgccaccca ggatctgggt 9960
tcctccttg gggggcccca gggaacaagt cacctgtct tgcataagg gagccctca 10020
gctatgtga gaaggttctg ctctgccct tcctcctct aggtgtcag ctctccagc 10080
ccactagtca gatgtgagg tccccagac cctgggcagg gtcattctg tccactgacc 10140
tttgggatgg gagatgagct ctggcccct gagagtcaa gggctggtgt ggtgaaaccc 10200
gcacagggtg gaagtgggca tccctgtccc aggggagccc ccagggactc tggcactgg 10260
gcttgccgt ggcatgtca gtcctccagc acttactgac accagcatct actgacacca 10320
acatttaca acaccgacat tgaccgacac cgacatttac cgactgac atttaccac 10380
actgtttacc aacctgaca tctactgaca ctggcatcta ccaactga catttaccga 10440
cactgacatt taccaacct atttaccac actgacatct actgacattg gcatctacca 10500

acaccaacat ttaccgacac caacatttac caacactgaa atttaccgac accgacattt 10560
accgacaccg ttaccaaca ccgacgttta ccgacaccga catttaccga cactgatatt 10620
taccaacact gacatctact gacgctggca tctactgaca ccgatgccag catctacca 10680
caccgacatt taccaacact gacatttacc aacactgaca ttaccgaca ttgacattta 10740
ctgacactga catctactga cactggcatc tactgacact gacgtttacc gacactagca 10800
tctactgaca ctgacattta ccaacaccag catctacca caccgaeatt taccaacact 10860
gacatttact gacactgata tctactgaca ctggcatcta ctgacacca catttacca 10920
caccgcatc taccaacacc gacatttacc aacaccagca ttaccaaca ccgatgttta 10980
ccaacgccga cggttaccga cgccagcatc taccaacact gacatttacc gacaccgaca 11040
tttaccgaca ctgacattta ctgacactga catctactga tactggcatc taccgacact 11100
gatatttacc aacgccagca tctactgaca ctgatgttta ccaacaccga catttaccgag 11160
caccgacatt tactgacacc aatatttact gacatcaaca ttagccatg tgatgggggc 11220
cggcttgggg gcaggccttg ctcttggcac tggggatgct gcagagacca gacagactca 11280
tggggatcag gacttctgct tcttctccag cctcatgtac tggacagact ggggagagaa 11340
ccctaaaatc gagtgtgcca acttggatgg gcaggagcgg cgtgtgctgg tcaatgcctc 11400
cctcgggtgg cccaacggcc tggccctgga cctgcaggag gggaagctct actggggaga 11460
cgccaagaca gacaagatcg aggtgaggct cctgtggaca tgtttgatcc aggaggccag 11520
gccagccac cccctgcagc cagatgtacg tattggcgag gcaccgatgg gtgcctgtgc 11580
tctgctattt ggccacatgg aatgcttgag aaaatagtta caatacttc tgacaaaac 11640
gccttgagag ggtagcgcta tacaacgtcc tgtggttacg taagatgta tcattcgcc 11700
aggtgcctgt agacacagct acttggagac tgaggtggga ggatcgctgg agtccaagag 11760
tttgaggcca gcccgggcaa aggggacaca ggaatcctct gcactgctt tgccacttac 11820
tgtgagattt aaattatttc acaatacaaa attaagacaa aaagttaatc acatatccac 11880

tgccctgctt aagacagaaa acatgggtgt tgtgaagcc agaggcagct gctggcctga 11940
gttgggtgat tggttcctaa gcagtgaag gcagtttgt ttccatag atgtctgttc 12000
tccctttgct ggggtgcagcc tcgccctgct gctgtggtcg gggttcagt gcctcgtccc 12060
gtggacgcag cctcgccctg ccgctgtggt cgggtttcag tggcctcgtc ccgtggacgc 12120
agcctcgcgc tgcgctgtg gtcgggttc agtggcctcg tcccgtagc gcagcctcgc 12180
cctgccgtg tggtcgggtt tcagtggcct cgtcccgtgg acgcagcctc gccctgccgc 12240
tgtggtcggg ttccagtggc ctgctcctgt ggacgcagcc tcgccctgcc gctgtggtcg 12300
gggttcagt gcctcgtccc atgggcgtgc ttggcagct ttgtctcac ctgtggagcc 12360
tctctgagc tttttgtt gtgtttgtt ttgtttgat ttgtttgat tgtttgtt 12420
tgtgtcgtt gttgtgccc aggctggagt gcagtggcgc gatctcagct cactgaaacc 12480
tctgcctct tgggttcagt ccattctct gcctcagcct cccacatagc tgggattaca 12540
agtgcgcgcc accacgcctg gctaaattt gtatttttag tagacagggg gttcaccat 12600
gttggtcagg ctggtctgga actcctggtc tcacatgac cactgcctc ggctcccaa 12660
agtgttggga ttacaggcgt gagccaccgc gccagccct ctgttgagca tttttgagg 12720
ttctcttgt gccagtata gtacatgtg tccccatgc accatcgtca ccattgagg 12780
tgacatttgt gcctctctc ggggttgatg cctccctctg ttccagcaa ctctgaagg 12840
atttcctga gctgcatcag tccttgtga cgtcaccatc ggggtcacct ttgtctct 12900
cagggctccc aggggaggcc cgaatcaggc agcttcagg gcagggcagg atggagaaca 12960
cgagtgtgtg tctgtgtgc aggattcag accctgcttc tgagcgggag gagttcagc 13020
acctcaggg tggggaacct agggatggg gaggtgagt ggacgccct cccacgaaa 13080
ccctaggagc tgcaggtgtg gccattcct gctggagctc ctgtaaatg tttgtttt 13140
ggcaaggccc atgtttcgg gccgctgagg atgattgcc ttcacgcatc cccgctaccc 13200
gtgggagcag gtcagggact cgcgtgtctg tggcacacca ggctgtgac aggcgttgtt 13260

ccatgtactg tctcagcagt ggttttcttg agacagggctc tcgctcgctc acccaggcga 13320
gagtgacgtg gcgcaatcac ggctcgctgt agcctcaatc tccctgggct caggtgatcc 13380
tcttgectca ccctctgagt agctgggact acagacacat accaccacac ccagctagtt 13440
tttgtgtatt ttttgtgggg ggagatgggg ttctgctgtg gtgcccaagc tgatctcaaa 13500
ctctgagggc acaagcgatc cacctgcctc ggctcccaa agtgctggga tgacaggcat 13560
cagccgtcac acgcagctca atgattttat tgtggtaaaa taaacatagc aaaaattga 13620
tgattttaac cattttaaag tgaacagttc aggctgggcg tggaggctta tgcttgaat 13680
cccagtactt tgagaggctg aggtgggcag atcacctgag gtcaggagtt tgagaccagc 13740
ctggccaaca tgatgaaatc cagtctctac taaaaataca aaaattagcc gggcatgggtg 13800
gcagggtgcct gtaatcccag ctactgggga ggctgaggca ggagaatcgc ttgagcccgg 13860
gagggtggagg ttgcagtgat ctgagatcat gccactgcac tccaatctgt gtgacagagc 13920
aagactctgt ctgaaaaat aaataaataa aaaaaattt aaaaagtga caattcaggg 13980
catttagtat gaggacaatg tgggtgcaggt atctctgcta ctatctactt ctagaacact 14040
ttctctgcc ctgaaggaaa cccatgccc accggcactc acgcccattc tcccctctct 14100
cccagcctct gtcaaccact aatctacttt ctgtctctgg ggggtcactt ctctggagc 14160
tttgtgtga ctggaatcct gcaatatgtg gtcctgcgt gtggcttctt tccatagcat 14220
tgtgtttcc agattcacc acacattgtc gcacgttatc agaattctcat tctgactgg 14280
gtgcagtggg ttaggcctgt aatcctaaca ttctgggagg ccaaggcggg acgatcactt 14340
gaggcaggag ttgagacca gcctggccag cctagcaaga cccagctac caaaaaattt 14400
taaaagttaa ctgaacgtgg tgggtgggg cacttgtgtg tcccagctac ctgggaggct 14460
gagggtggag gatcgcttaa gccaggagg tcaaggctgc agtgagctat gatcgacca 14520
ctgcactcca gcctggacaa cagagcaaga ccctgtctga aaaaaaaaaa aaaaaaaaaa 14580
gttctttct ttttgtggct ggatgacatc ccattgtatg gccacagcac atttgtttg 14640

tctgtttatc ggggtggggg cagtgggttc caccttttgt ctctgtgaa taatgctgct 14700
gtgaacattt gaattcaagt tttgtttga acacctgttg tgaattattt ggatatatgt 14760
gtaggggtag gattgctgag tcctatggta atgtaggtt tgacttactg aggaaccatt 14820
aaactgtttt caacagtggc tgcgccgttc tgcattccca ccggcagtgt gtgagggttc 14880
tgactttacc tctcacaaa cgcttctttt ccatttaaaa aaatattcag ccagggtgctc 14940
tggctcacgc ctgtaatccc agcactttgg gaggccgtgg cgggcggatc acctgaggtc 15000
aggagttcga gacgagcctg gccaacatgg tgaacccca tctctacca aaatataaaa 15060
attagccggg tgtggcagcg ggcgcctgta atcccagcta ctgggaggc tgaggcagga 15120
gaatcacttg aaccgggag gcagagggtg cagtgagcca agatcgcgcc actacactcc 15180
agcctgggtg acaagagtga aactccatct aaaataaaac aaaaataaaa ataaataaaa 15240
atttattaaa acattcatca cagccagcct agtgggtgct ccatgtggct ttgcctcgca 15300
ttccctgat aactaggatg ctgagcgtct tgtcccaggc ttgccacacc tcagcactt 15360
gagatacgtc gcacagtccc catttgcgaa cgagaaatga ggttaggga acagcagctg 15420
tgtcatgtca cacagcgagc aggggggtctc tgagccgtct gacccacag ccgaccaagc 15480
tccaatcctt accgcctcct agtgtgtgg atgtagccca ggggtgctcc acattttca 15540
gatgagaaca ccgaagctca aaacaggagc gttttgtcca cattggatac acgatgtctg 15600
tggtttggtc ctgaagtcac ttatatctc agtgggtccag actggagtag gacagggggt 15660
tctggggaat ggggaagggtg tctcaggtga aaggaaggaa ttccagattc tccatactgt 15720
ccttgggaag ttagaagact cagagggtct ggcaaagtca gacaaagca gagaaatgca 15780
gtcaggagga agcggagctg tccaggaaca ggggggtcgc aggagctcac cccaggaac 15840
tacacttgct ggggccttcg tgcacaatg acgtgagcac tgcgtgttga ttaccactt 15900
tttttttt ttgaggtgg agtctcgctc tctgcccag tctggagtgc agtggcacga 15960
tctcggtca ctgcaagctc tgcctcccg gttcatgcca ttctctgcc tcagcctccc 16020

gcgtagctgg gactacaggc gcctgccacc gcgcccggct aatitttgta ttttagtag 16080
agatgggatt tcactacatt agccaggatg gtctcgatct cctgacctca tgatccgccc 16140
gtctcggcct cccaaagtgc tgggattaca ggcgtgagcc accgcgcccg gcccgatttc 16200
ccactttaag aatctgtctg tacatcctca aagccctata cacagtgcctg ggttgctata 16260
gggaatatga ggcttacagg ccatggtgct ggacacacag aaggacgga ggtcaggagg 16320
tagaagggcg gagagaggga acaggcggag gtcacatcct tggcttcaa aatgggccag 16380
ggagagacac cctctgagca tggtaggaca ggaaagcaag attggaacac attgagagca 16440
accgaggtgg ctgggcgtgg tggcttacgc ctgtaatccc aacactttgg aaagctgagg 16500
tgggtggatt gcttgaggcc aggagtcaa gaccagcctg gccaacatgg tgagaccccg 16560
tcttactaa atatacaaaa attagccagg cgtgatggtg catacctgta atcccagctg 16620
cttgggaggc tgaggcagga gaattgctta aacctgggag gcggagggtg cagtgagccg 16680
agatcccgc actgcactcc agcctgggcc acagagtgag actccatctc aaaaaaaaaa 16740
aaaaaaaaa taaaaagacc aaccgaggaa ttgaagtggg ggggcgctac agtagcagaa 16800
gggggatcgt ggagcaggcc accctgtggt catgcactgg aagctcatta cctgacgatt 16860
tgagctcat cactgggggc ctaaggagaa tagatactga aggatgagga gtgatggcgc 16920
ggggcacggg tgtctttggt ggccagaact tggggactgc tggggtgcct cactgcaggc 16980
cttctcagcg cctttatat gcttacacag gctgttcta agagggggat acattgcata 17040
agcgtttca gactacctca tcatgggtcc cttctttac cctctgtggc cctggtggcg 17100
cactctctgg gaaggtgcag gtggatgcc agaccgccc tgccatccac ctgcacgtcc 17160
agagctgact tagcctcgag attgctgctg gcacctcctg ccccgggaca cctcggatgt 17220
gcccgtggag atgctggctc tgtgtttct gctggagttt ggtgcgtctt ttcctcctgc 17280
aagtggccac cgctctggg tatgtcctca ggcttctgcg agtcatggct gcttctcagg 17340
tccttgccca gcgccaggag caaacctcc tggcacttg ttcaggggtg gatgcgccag 17400

tgttcctgct gtggaccgcc atctacatg agggctctgg gcctgcaggc tcgttcagga 17460
aacacccgct gagtatgcag tgtgtgccag ctgtgtccca ggcaatggcg gggacagtgg 17520
ctgtctgctgg gggtgtggg gcttctgggg actctgggga cagctgaggt gcaaggagcc 17580
acggctcctt gaggatgcag ttggactcca ggtggaaggg atggttgggg gaggtataaa 17640
tggggtcagg gaggagacac atttgaaca atgggaacat ttttaagatg ctatgtcggg 17700
aggcaacaag gtggccaacc caggtgctga ggagcccaca ccagccctgg acgtgtttg 17760
ccgctcacct ttgctgggga gtgggtgggag agaggattcc gttccacgtg gtggtgtgcg 17820
cagctgggct gtgtggagct gggcgctagg aggaaggtgc ttctgcggg gctagccggg 17880
ctctgccttt gaacacaatc aggtccagg tttcagcat ccagtgcag agaggacttc 17940
acgggcagct gtggctgac ccttgatgaa ttgggagaag aacaaaggtc tatgaaatga 18000
ggtttcagtg agatggcatt agagacgcc acaacagatt tacagagtgg agcggagacg 18060
gcggatgggt ctgggaggcc cctcctgctg gccttgactg tgacagctgt cctgggaatc 18120
agcttcagg ccgccccagc agcctgactg acacacacag gggtttagc cccatcctgc 18180
gaccagctgt tgccatcatc agtgacagct gggagtggcg gtggttcag ccctgggcac 18240
cctccccacc tctgggggcc caccagggc agtcctgaca cctacaggtt gcttgagcc 18300
gcatccgagt cctgccccac cacgtgtgaa gcccgagtgg tctggggctg aggtcccctg 18360
attgcatccc cacttccctt ctgttcaca tagctgctc ttctaccgt tttccagcc 18420
tcctgggcta ggaattccag tgtgtgctg gcttgcccc aggacacctc ctagccctc 18480
ttctgagtc tagagccccg ggggttgaa gtctggccc ctgggacacc tgcagccaca 18540
ctcagcttct cctgtgagcc tccagcatgt cccctcagga ccaagccctc acgttctgc 18600
ctccccgcc acctgggctc agccaggga aggcctggct gggagcgtct cccctctgcc 18660
ctgcccttct cccctctac cctgccctc tctcctctgc cccgcatgg cttttatc 18720
ctgtgccaca agacatggct gtgtgtgaaa gtggcagggt ctggcatctc tgtgggtctc 18780

tgaggccac gctccagtgc cactcttccc acccgctggc cgtgccctca tgctggaggg 18840
acagcccagc cctctcccga accccagccc catgtgcccga gctgcccccg gccctctccc 18900
ctggaagccg gggtcactcc agccgtatgc catgggtggg acatcctgct tccttggcct 18960
tccagggaag gtctcttcc caaatggcga cacctgggcc ctgcctggag gctggaagct 19020
gtggcccttg tatgcccctc caggggtctgt gcgctcggtt ggcccagatt cccatcaccg 19080
tcatcatcac catcatcatt gtcatttcgc ttgtctgtga gccggcctgg tctcccagag 19140
cagagaccct ctgagggtcca gcctgagttg gggctcctgt gctgaccct gacggggact 19200
caggacgtac caggtctggg tcaggagtga ccccaaac tcgtgccctt tgacaggcac 19260
ccctgacttt tgctaagtgg gtggaggtga catcacttac agcgggagtg atgggacagg 19320
gtctgttggc tgcactgtgc tcccaggat ctggggagag gctatatccc tgggcttgg 19380
cactgcagag ctgtgtgtgt ttgtgtgtgt gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt 19440
gtgtgtgttt gcgtgcgcgc acatgtgtat aagatctttt ttattacat gaagcaagat 19500
aactgttgc gtttctttt gggttttgtg ttcaacagag tggggtactt cttccctcag 19560
acaacagaac tctccccttt aaacacgtgc tgcagaggg tgggtcttgg gctcatgtct 19620
gtttgcacag ccgagtcaga ggaaacacag ggttcttcat aaaaacactg cacagcaggc 19680
gactgtccag agtcagcctg caggacggca gcagccctgc cctcagagc acagctaggg 19740
tgggctgctt tgggatctcc cgtcattccc tccagctgg cagccggcgg ccggcccatt 19800
ccttgggtgt ctggtcaggg gggcgtgcgc ctgctctgct caccctggga atgggacaga 19860
agctggcagc tcggagagga cagggtgga ccttgggtg gcctctggct ggaccatctc 19920
attgtcctca gacacagcct ctgggtcta gtttcattc ctgaaaaa agtcacaga 19980
actagagcag gagtgcagag ctacggcccc cgggccagat ccagccctgc cacctgtttt 20040
cacaccatgc tcaagctgag tgggttttac atttttaat tacttgaaaa aaaaaagcc 20100
aaaggagggt tcatgacca tgaataatt atggaattca aaaaaaaaa attatatga 20160

attcaaatTT cagtgtccat aaataatttc ttgagacagg gtctcgctct gtcacccagg 20220
ctggagtgca gtgctatggc atggctcgct gtacccttga cctcccaggc tcaagcgatc 20280
ctcctgtctc agcctcctga gtagctggga ctacgggtgt gtgccaccaa gcccggctaa 20340
tttttttta attttagtaa agacagggtc ttctatgtt gcccaggctt ttctggaact 20400
ccatcttggc ctcccaaagt gctgggatta caggctcgag ccacggagcc cagcctgttt 20460
ttgtttttc actgataaag ttttccggg tgtggtagt gtgcctcta gcgattggg 20520
aggctgaggt gggaggatcg ctaagccca ggagttgag gctgggctca agtgatcagg 20580
aggTgaacta tgatcatgtc atgcatcc agcctgggtg acagagcaag aacatatctc 20640
ttaaaaatat atatttaaaa agtattgggt gtggtggctc acgcctgtgg tcccagctac 20700
ttaggcattc gaggTgggag gatggctga gcccaggagt ttgaggtgc agcgagccaa 20760
gatcgtgtca ctacactcta gcctgggtga cagagcccag accctgcctc tttaaaaaa 20820
aaaacaaaa aacatgtatt ggaacacagc catgcctgtt cagtcacgtg ctctccatgc 20880
tgctttctgc tccagagacc ctatggcct gaaagctgaa aatattttct atcctttaca 20940
aaaaagtTtg ctgaccttg tcttgaaaa tcatctccc aagtctctt ccggcactgg 21000
cgTtctggg tgcctaaat ttggcccctg ttatttctga actctgttt ggctctgttc 21060
cctcccagga gccaggacag gcacgttctc tgcatttgt cccctgacgc ccagaggctt 21120
ggctcggctc aggcattctt ggaaatatct ggctccagga aaggcagagg cctcctgagt 21180
cggcccagag ggaacctgcc ccaggtctgg gggaggcctg acccagcaga gtggcttttg 21240
ccgatgggtt gggccgggtc agatgtgctg aaagtgtcc tcagaaggcc actttgggat 21300
tcttctctc agtattagag caactgagag ctgctcattg caagcctgat gtttcccag 21360
ttggccgggt ccaccgggtg ccctgggatt ctgggatctg ggtggaaagt agggggcttg 21420
ggggagtgtc ctgggttctg gaatccaggt ggcaagtgt gaggttcagg gattggcttc 21480
tgagccacca taggggtctc tgtgggaggc tctgccatc caggagattc cgcaggccct 21540

gccggccag agccagcgtc ttgcgcttgc cgaggctaca gccagcccca gccgggtgga 21600
acagcccgtc gcctcctctc actttgtttt ggggccacct gggagtgtgg agcaagggtg 21660
gagagggagg aagtggctgc cggccgctgc ccagcaccct tgtttgcctt gggccctctg 21720
tgggctcctt ttattgtctc ttcaatgaag ccagggaat ggacttcctt gcctcacttc 21780
agttcaacat gtctggaagt ttgtattaa aattaagaaa gtgtggaaat agagcaagaa 21840
gagaaaaatc tctccaagag ataatagtga cctctgagct gggcgcggtg gctcacgcct 21900
gtaaatccca gtactttggg aggctgaggc gggcagatca cctgaggtcg ggagtgtgtg 21960
accggcctga ccaagatgga gaaaccccgct ctctactaaa aataaataaa taaataaata 22020
aataaataca aaattagcca ggcatggttg cgcctgccta taatcccagc taaggcagga 22080
gaatcgcttg aacctgggag gcaaaggttg cagttagcca agatcacgcc attgcactct 22140
agtctgggca acaagagtga aactccgtct caaaaaaat aaataataaa aaaataaaaa 22200
tagtgacctc tggccagggtg tggcagctca taccgtaat ccagcactt tggaaggaag 22260
gccgagatgg gcagattgct ttagcacagg agtttgagac cagcctggcc aacatggttg 22320
aaccctatct ctacaaaaat agaataaaat ttaagaggta atagtacact ttggttagat 22380
cgaaacctgg attgctttct ttctctaaat gctgattctt ttctttgttg tgtttgtgtt 22440
ctgtgccgat gtccctccc cagccctgtt attgtgagtg gaagaagggg aaagggttcg 22500
cccgtactg tgagccctc ctctcacgct ggggtgcctt ggagaagcct gcacttcttc 22560
attgtacgcc agggctgggt cctcccttg agtggttctg tgctgctggg atggggccaa 22620
cccctcagat gttttctgag tgtcacacac aggtgtgtgc attcatggcc ttgcgtgtc 22680
ttcctgttgt ggaggcaaaa atgtgaagaa ccctagatga tttgggacc agggctccat 22740
cacctgctgt tcattgcaca ccggagcatc caggcatggg tggagagctc agacttcag 22800
gcacggctgc aggggctggt ctaacatgt tccgcccgc ctgctctca gaaccgctg 22860
ttgggagctg ttatcatgat accatacctg ggccctgggc tatccgattc tgacttaatt 22920

gctccagggtt ggggccaggc cgttgtttgc tgtttgttg tttctctgt gacgttagcc 22980
actgggctaa tctgagcccc tcagttacag gtggagaaac tgagacccat ggggggtgcaa 23040
ggacttgccg aggaccaga gccccttggg ggcagagctg aggcggggcc tggctttggg 23100
tcccagagct tccagtcccc tccccgtct cctaacagct tttttttg agacaagatc 23160
tcacctgtc acccaggctg gagtgcaatg gcatgatctc ggctcactgc aatcttcgt 23220
agctgcgttc cagcgattct cctgcctcag cctcccgagc agctgggatt acagggtgtg 23280
gccgcatgc ccagctcgtt ttttttga ctttagtag agatagggtt tcacatgtt 23340
ggccaggctg atctgaact cctgacctca aatgatccgc ctgcctcggc ctcccaaagt 23400
gctaggatta caggctggga tcacactgtg cctggcccta gcagctttgt cctgtgcat 23460
ccaacaacag atgaccgaag tctttgttc ttaacatgca ttccatctgc cttacagttt 23520
tgccacctgc aaaacagagg actgtcgtc tttctggtaa gctggaaatg taatctggta 23580
gcaggaggcc tgtggaagct tgccttaat ggccttgtgt ctcttcac ctgtcctgag 23640
agccggagaa ctgggatgtt gcacctact caaccttct gttacatac agttctgcag 23700
gctcatgat catcagaacc acgtcctatc tcacgcggct gtatgctcc gttggttcag 23760
gtgttttac ctgacagta tttctcctc ggtggcttt gcggtggtt ctttaatca 23820
gcattgactc ttaagaaaa atatttagct gctacatctc agaggagaca gggtggaag 23880
catctgagac ctgcaggctc agacttagaa ccagaagtgc cctcagagt catccggccc 23940
tgacccagcg ggaaatgagt tcacagagaa gcgggagaaac ttgccccag gccctgccgt 24000
tgctcataac tgccccaggt cttacattt gctccaggtc ctgccccagg ccctgcagtt 24060
gctcataact gccccaggct cttatattg ctccaggctc tgccccaggt cctgcagttg 24120
ctctgtgtgg tgggtgtgat ctggagccct ccgcccattg ctgcacctgg ggcaggcatt 24180
gctaattgat cccaggactc ctctctcgg agcacgccct ggttctccag gcagccgctg 24240
cctgtcagcc tgcagtgtt cgggagagga cacctgctg cctggtctgt tccaaatctt 24300

tggagagagc agcatccagg tgcggcaggg acaggcctgg ggctcgcggg cagggactct 25740
gtgtcctgcc ggggtccac actgcacctg cttgtcagag gcactcagtc aatctttgct 25800
gatgaaggat gagaggacag aggacgtgat gcttgctgct gcattgcctg cagtcctggg 25860
tgagatgccc gggttgactc tgctgcccgt cgggtggatg tgatgicaga tccccggctt 25920
taaaatacga gggagctggg aattgaggga gcaggttggg gcagaaagca cagccccgtg 25980
gaagcctgga gctgaggcag tgtgggcgac ccctggagca gtgagtgcctt ccttcacggc 26040
cttcacgca ccctgcagtc ctcatgtagg ggatgccatc catgaattta gtttcccag 26100
cctccttaa aaacgcgttc atgctggggc cggggcagtg cagtggctca catctgaaat 26160
cccaccactt tgggaggccg aggcgggtgg atcatgaggt caggagatcg agaccatcct 26220
ggctaacaag gtgaaacccc gtcttacta aaaatacaa aaattagccg ggtgcggtgg 26280
cgggcgcctg tagtccagc tactcgggag gctgaggcag gagaatggcg tgaacccggg 26340
aagcggagct tgcagtgagc cgagattgag ccactgcagt ccgcagtccg gcctgggcga 26400
cagagcgaga ctccgtctca aaaaaaaaaa aaaaagtaca aaaaaaaaaa aattagtctg 26460
ggtgtggtat cacgcgccta taatctact actcgagagg ctgaggcgga gaattgcttg 26520
aaccaggag gtagagggtg tagtgagccc gtatcgtacc actgccctcc acctgggcaa 26580
tagagcgaga ctctgtctca aaaagaaaaa aaaaaaaga acatttatgc cagggtgtgt 26640
ggctcatgcc tgaaatccca gaactttgga agactgaggc aggaggatca cttgagccca 26700
gaaatttgag agtgtcttc ctgggcaaca tagagagacc tcattctctac cagaaaaaa 26760
aaaattagcc cggcatggtg gcataccct gtgtgccag ctacttaggg ggctgacgtg 26820
gcaggatcac ctgagtctgg aggcagaggt tgaagtgagc tgagatcatg ccactgcact 26880
ccagcctggg tgacagacag agaccctgtc taaaaaaaaa aaaaaaaaaa aagcatttac 26940
tatccaccat ggaagggtgag actgacctgt gagtgtgtgt tcaagaaca aaaaataaac 27000
cccagagata agacaaaagg gtgcctccat ggggggtgtga tttaaagctg agaaattggg 27060

cttcttcccc ctccccctctc accccgtggt ttgctaaagg agatgggaaa aaggattctt 27120
tttttggtcg aaatatatta cactaaatta aagccaattt taacagcact ttggtgatg 27180
agtgaatta acagactggc caaaaataaa cgaacggtct gtactatgtg aaaaagaggc 27240
agctttggcc atgctgggcc aatgtgagtt ttcagggttg ctgggaatgt ctgtgaatcg 27300
gaggaagggc ctagctggga ctctcaggag ccaaggccct gaggggcaac tgcctggtc 27360
cctgccctga ggcgttact gctttcttc tgggccagat cacaggcccg gaggctggac 27420
cactgggctg gcactctgc cgagctgctc cctgactcc tgaccatgct ccttcagca 27480
gccttgctgc actttagttt ccttgaatga aaaatgggga tgagaatagc tctacctcc 27540
aaggagaatg gagtgagtc ggacaggatga ctccctggga ccagtgcctg ggcctgaca 27600
aggtccagtc agagcccgca ctgctgttac tgataccctt ggctgtacca ggggagaact 27660
tggttgccat tgccaggtgt tctccacca cccccactac tgcctgtt tgatgtgtg 27720
cggaataaaa gctgtgcaca ttggagcttt tggcacatcc tggcttcag gtgaaagggtg 27780
cgtgtgtgtt tgagggttta gcctggccaa cccagccatg aggtcggacc tgacctgggg 27840
gtgagtcctg agtcggcac ccctgagctg tgtggctcac ggcagcattc attgtgtggc 27900
ttgggccgca ccccttccc tgctgggctg ttgatgtta gactggagcc tctgtgttcg 27960
ctccaggaa ccaaccgtg tgcggacagg aacggggggt gcagccacct gtgcttctgc 28020
acacccacg caaccggtg tggctgcccc atcggcctgg agctgctgag tgacatgaag 28080
acctgcatg tgcctgaggc ctcttggtc ttcaccagca gagccgcat ccacaggatc 28140
tccctcgaga ccaataacaa cgacgtggcc atcccgtca cgggcgtcaa ggaggcctca 28200
gccctggact ttgatgtgc caacaaccac atctactgga cagacgtcag cctgaaggta 28260
gcgtgggcca gaacgtgcac acaggcagcc ttatgggaa aacctgcct ctgttctgc 28320
ctcaaaggct tcagacatt ttctaaagc actatcgtat ttattgaac gcagtcaag 28380
ctaataaat atgagcaagc ctatttaaaa aaaaaaaga tgattataat gagcaagtcc 28440

agcctattct actgtttgta ttacatagct ttaaagatt tttatgact ttaagtcaca 29880
agggttcttt gtagaaaaa atatatatat aggaaagtat aaaaagaaag taaaattgt 29940
ccataacctc tccagccaga gacgaccgtt gctgacacct cagcatattg cctttaagtc 30000
tttttctct aagatagcat ttctttcat cacagtcata tgctacgcag aattctgtat 30060
cctgattttt tcacttgaca ttacaacagg tattgatgg cgcttgaca aactcttgg 30120
cacaatcttt taaatgtatg aaatactcca ctgcacagat gtttgctttt aggettaact 30180
gttcttttat ttgctgtg ctggttacag ccgggcacag tggctcatgc ctgtaacac 30240
aacatttga gagggtgagg caggaggatc acttgagccc agaagttga gaccggcctg 30300
ggcaacatag tgagacccca tctctacaaa aaactttttt aataagtcgg gcgtagtggt 30360
gcatagctgt agtcccagcc accaaggagg ctgagttggg aggattgctt gagccccagg 30420
aggttgatgc tgcagtgcac tgagattact ccaactgtact ccaacctgag cgacagagca 30480
agacttgtct ggggaaaaaa aaaaaaaaaa tatatatata tatatatata tatatacata 30540
tatacataca cgcacacaca cataatataa aaatatatat ttataaatat ataatatata 30600
atataaaaat atatatttat aaataaaatt tataaattat atttataagt aaatatataa 30660
tatataatat aaaaatatat attatataat atataataa atatataata taaaatatata 30720
tatttataaa taatatataa tacatactta taagtatata tttaaatat atgtaatgta 30780
tattttttaa tgtatgatat ataatatata ttataaata cacatttata ttattttata 30840
taaaatatat ataaaatctc caagttgctt ttccaaaaa ggtgtcttgc tgcatttcaa 30900
acattcattt aaaaacttga atgctgggtga tctgggtccag aatgtgttca gtagctgctg 30960
ccagtggcca agcatctcgg gagatgtcta caaaacacgc tggttctggc ctggcgtggt 31020
ggctcacgcc tgtaatctca gcactttggg aggctgaggg aggtggatca actgaggtct 31080
ggatttcgag accagccttg ccagcttggt gaaaccccat ctctactaat aatacaaaaa 31140
aattagccag gcgtggtggc atgtgcctgt aatcccacct actggggagg ctaaggctgg 31200

ggtagacaca cataagggct tttgtgaaat gcttgtgtga atgtgaaata ttgtgtgcc 28500
gttagagcttg acttcagaca cccacccac tccctgtcg gtgcccgttt gctcagcaga 28560
ctcttcttc atttatagt caaatgtaaa catccaggac aaatacagga agacttttt 28620
tttttttt tgagacagag tcttactctg ttgccaggc tggagtaccg tagcgtgagc 28680
tcagctcact gcaacctccg cctcccagg tcaagcgatt ctctgcctc agcctcctga 28740
gtagctggga ctacagacat gcaccaccac acccagctaa tttttttat atttttagta 28800
gagacagggt ttcacatgt tggccaggct ggtcttgaac tcctgacctc aggggaacag 28860
acgggggttg cctcccaaag ggcgggaaata acaggggtga gccaccgtc ccggcctagg 28920
aaaactttt gccttctaaa gaagagtta gcaaactagt ctgtgggctg gccttctgat 28980
tctgtaaaga aagtttgatt ggtggctggg tgcgggtggc cacacctgta atcccagcac 29040
tttgggaggc cgaggtgggc agatcacctg aggtcgggag ttcgagacca gcctcacaa 29100
cgtggagaaa ccccgctct actaaaaata caaaaaaaaaa attaacggg catggcggcg 29160
cctgcctgta atcgagcta ctcaggaggc tgaagcagga gaattgcttg aacctgggag 29220
gcggagggtg tggtagctg agatggcacc attgcactcc agcctgggca acaaaagtga 29280
aactccgtct cagaaaaaaaa aaagtttgat tgggtgaacc aaagcgcat tgttatgga 29340
ttgtctgtg cagctttgt tctgccgaga tgagttgtga cagatctga tgggctctaa 29400
agcctaaaac atgtgccatc cgcccctta cagaaaaagt gtgctgacct ctgttctaaa 29460
gtattggaca actacaatgt ttgtcattt attattctat gatttgttt ctgcttttg 29520
ttgtgtgtg ttgtgttgag atagggttc cctctgtcac tcaggctgga gtgcagtgg 29580
gtaatttcag ctactgcag cctcgacctc ctgggctcta gtatcctct catctagcc 29640
tccctagtag ctgggactac aggcacacac caccactcct ggctgattt tttttttt 29700
tttttttt gtggagacag ggtttccgca tgtgcccag gctggttca aactcctagg 29760
ctcaaacacc cacctcagcc tcccaaagt ctgggattac aggcgtgagc caccatgccc 29820

agaatcgctt gaacccaggg ggcagaggtt gcagtgagcc gagatcgac cattgcactc 31260
caggctgggc aagaagagcg aaactccgtc tcaaaaaaa aaaaaaagat gctggttcct 31320
aaaatgtggc ccttttcctc ctcacctgt gccagacat cagccgccc tcatgaacg 31380
ggagctcggg ggagcacgtg gtggagttg gccttgacta ccccgagggc atggccgttg 31440
actggatggg caagaacctc tactgggccc acactgggac caacagaatc gaagtggcgc 31500
ggctggacgg gcagttccgg caagtcccg tgtggaggga ctggacaac ccgaggtcgc 31560
tggccctgga tcccaccaag gggtaatgt ttgcctgtcc cgtgcgtcct tgtgttcacc 31620
tcgtatgaga cagtgcgggg gtgccaactg ggcaagggtg caggctgtcc gtgtggccct 31680
cagtgattag agctgtactg atgtcattag ccttgatgtt ggccaggact ggtagggccc 31740
tcagaggtca tggagttcct tcgtggagcg ggtgctgagg ctgtatcagg cacagtgtg 31800
gctgctttca cctgggccgt ctcaccgaag tgtccatgga gcctgcgtag ggtgggtatc 31860
tgtgtcgatt ttacagatgc agaaacaggc tcagagaaac cgagtgactt ccctaaggtc 31920
acatacccag ttagagcaga gctgggccag gaagtgtgt ctcaggctcc tgaccaggtc 31980
tccttgctt gcactcttc caaaaccatg atccagaact gactttgagg tccccggacc 32040
tcaggctcct ccgaaatggc ctcttgagg ctgctgagcc acagcttagg acccacctcg 32100
agaggcaa atgtcttgag ctgccaggcg tcctgggggc cctgccttgg gcacggggtt 32160
cagacaggcc ccagatgtgt ggggcgtct tctggacttg agttttctt tctgtgtgt 32220
ggacacagt ctcacccctt aaagcacctg tgatgtgtgc agcagcccaa tccctgcctg 32280
tcgcctgtt tgctaggga ggaaggaata ctcaggatg gcaggacaac agaaagaggt 32340
ccaggtttta gagcaagggc aggtcaaact tagaaaatc tggaatgagg atgtgcattt 32400
cctcttctgg atctgctaaa agaagaggga aggaggggct gctgggggag gagcccagag 32460
ccgagtttac atccggatcc cgcaaggcct cccctgccct gaggtcttgt tttgtgatgt 32520
gcttgtgtcc atcctgggtt ctgccgtgc cccaacatcc ggccaagctt aggtggatgt 32580

agtcccagct acttgggagg ctgaggcagg agaatggcgt gaacccggga ggcggagctt 42300
gcggtgagcc gagatcgctt cactgcactc gagcctgggc aacagagcaa gactccgtct 42360
cacgcaaac tctgtctcac gcaagactcc gtctcaaaaa aaaaaagagt tcagggttta 42420
tgaaactggc cagccgcgta aagtttgctg tgttgtttt gtgcccggga ggagtgtggc 42480
cagggtgtca cgtcacacag tacagtttc tcagatggg gttctccaga ctgctgtccc 42540
aaagtctgtt ttgcatctg gtccacacag acccaccctc cacggtgagc ctgattttgg 42600
ccagggtagc tggaaatttg ctgtcttc agcccggcag ctgtaccagt ccagggtcca 42660
cagctagtgg ctttaggaa ggaattgtt cagttggctt tgacacatgg cccctaggg 42720
tccacagctc tgtagtgatg tggatgtgt tatctacaaa gacacatgat ccttcgtgc 42780
cagatgaaag tgatgatgc ttgcagctg cccagcaagg ctgtgtgtgt gtgtgtgtgt 42840
gtgtgtgtgt gtgtgtgtgg tgtgtgtgt gtgtgtgtgt gtgtatggg gagggaggca 42900
ccctttccat ctgggggtgt gtgtgtgtgg ggtgtgtgt tgtgtgtgcg cgtgtgtgtg 42960
gtgtgtgtgt tgtgtgtgt tatgggggag gcacccttc catctgggc caagagactg 43020
ggcctgggga agacgttct tttatctac ttagagactt tgtttattt gtatttttt 43080
gagacagggt ctactctgt caccaggct ggggtatgt gatatgagca tagctcactg 43140
cagcctcggc ctcccaggct gaagcgatcc tccacctca gccttctgaa tagctgggac 43200
tgtaggcgtg cgtcaccata ctgagctatt gttttttt tttggttgt ttaattttt 43260
ttgatacaga tggagtctg ctatgtgcc cagactagtc tcaactcct gaactcaagt 43320
gattctccca cctcagttc ccgacattt gggatcacag gtgtgagcca ctgctgtctc 43380
cctgtttat taactgctga aagacctaga taaagaaagt ctgaaaagac ttactatcag 43440
agcaccatcc taagatgatt ccctctgact caatggagag ggaggggagc tttccttca 43500
ggcctgggtg gcaggagccc aggtgtcca ggccccattt gcccaggcc aatcactcg 43560
ggaacttga tgcagctgtc ttcagggtg acccaaagga acagatccc cgcaggcagt 43620

aggcttctgg gctgtcctct cctctacgt cagctcagta agagcccttc gaagggatgc 43680
tgtgtcggag gccccaaaag cccaggctca tccctgagat gcacagggtg ggctgggctt 43740
aggcagcgct cgagcatctc ctggacgggtg accccagaga gtgtggagac ggagagtcct 43800
tgagagtcac tgagagacgt ggctgccctg ccttcccaag aggggctctg agtcattccc 43860
cacactcacc tgccccctacc caccctcacc tggccccag cctcacctac cccacatct 43920
gtaccgatcc ctttaccgc accctccta cccaccctca cctcccctgt accctcacct 43980
ccccactca cccgccccctg caccctcacc tgtccccac cttcacctaa cccccacct 44040
cacctgccct cccctcacct ggcctccttc cgttggggaa ggggttgtaa ggggcggccc 44100
ccaaactgtc tgtcttggtg ccctgcagag aaaacagtac gtgagggccg cagtccaaaa 44160
gcttgagtcc tggaagggtg aggagacagg gatgtgttg gaaggggccc atggtcttg 44220
atcccttctc gactgtcaat ggggccttca tgggagcgcc agtctagtga tgcacagctg 44280
gggtcccggc ggggtggctga ggaggcctaa agtccgaggc ggcaagagct ctccagagg 44340
ctgtgtcct aatcgctctg gcatactcag gcgggcacgt agttaggagc tgattggaga 44400
ggagagaccc ccacaccaat actgggattt gactttcagg ctaaactga gaagtgtggc 44460
ctctgtgtc ctgccagagc tctccagcca gtgccaggg ctctccagcc agtgcccggg 44520
ggtctccacc agtgcccggg ggtctccgcc agtgccaggg gtctccgcca gtgccaggg 44580
gtctccgcca gtgctcagga gtcttggtt cttgtctta cagcccttg tttgacctc 44640
tctgagccaa ggccaaaacc cagacaggca gccccagac ctcagcatcg acatctacag 44700
ccggacactg ttctggacgt gcgaggccac caataccatc aacgtccaca ggctgagcgg 44760
ggaagccatg ggggtggtgc tgcgtgggga ccgcgacaag cccagggccca tctgtctaa 44820
cgcgagcgca gggtaggagg ccaacgggtg ggtgggggtg ctgcccgtcc aggcgtgccc 44880
gccgtgtctt ctgccgaatg ccagcctctc acaggctggg gagactttcc accctgggga 44940
tccaatgggt ggctttccag ggtcccaaaa gcaaacacag gctctttcac agcccctcca 45000

ggaaagcaga aagccccaag ggctggaagg gaagggggag ctctgtgag aggttacaag 45060
gcagcgctgg ccgacgggag ttgcagttga taggttttgt atcatccttg ttaaacttga 45120
accctgtgca gaaatccctt ccacggcatg ggggctgcct gttgactcgc tctgttcca 45180
ccacagggag ctctgggct tcttctccc agaggcccc gacgtccca cctgttggtc 45240
gtcagagctt ctggttggtg ggaaggcacc caggaccttg aggtctccag agagaaaagc 45300
cagggaaaga gggagaccga aacctatgtg acatgaaact caggctcaa actgagcacg 45360
ggaacgtttg gggacaggag cgcgatggcc ttctcagat agctgggggg ctggcatgaa 45420
gacgggagct acagccagca caggctctgg gccgggagcc cagagattga gccctgactc 45480
tgtcacttac tggccacgtg accttgggcg ggtggcatag cctcttgag actcagtttc 45540
ctcatggta ggagtgcagg ccacagtgtg gcggcctctg cagcacacgg ggggctcgtt 45600
gggcggaagc cccgggtcta taaggcggct gtgcaggagc cagccgagct ggtctccaa 45660
cagccagggc tccggggctc ttagcagctg tggggggcct gcacctgtt cccatggctg 45720
ctgtcagaaa ttaccagaag ccaggtggct gagagtaatg gacactgtt ctctcacagt 45780
tctgagggc tgaagcccga gatcagggtg tgggcagggc cctgcgcct ctgaaggctc 45840
tgaggggaacc ttgggcttc tgggtggctc aggcaccct tgacttgtg tctgtcact 45900
ccagtctctc tgtctggctg cacatggcgt ggcctcttct gtaccattga aggacacttc 45960
agttggattt agggcctacc ctacccatt gtggtcgtat ctgtacctt catgacattt 46020
gtaaagaccc tgcttcaaaa taagctcaca ttctgaggtt ctggggtgag cgggaatttg 46080
gagagcattg ttcaactagt atagaatgtg acctgtcagc ctgggcagc cctgagaggc 46140
aggggcttcc cacagcccag ctgggtgccc tgggctccgt gctgtccgag gagacgcat 46200
ccccacacc gtccttcacc cgccaccctc ccgcaggtag ctgtacttca ccaacatgca 46260
ggaccgggca gccaaagatc aacgcgcagc cctggacggc accgagcgcg aggtcctctt 46320
caccaccggc ctcatccgcc ctgtggcctt ggtggtggac aacacactgg gcaagctgtt 46380

ctgggtggac gcggacctga agcgattga gagctgtgac ctgtcaggta cgcgccccgg 46440
ggcctgccct aaccgcagac acccggcctt cattgtcagt aatggcagca gctgccacat 46500
tgtccgagac ctgccgtgag cccagtgccg cgccaggggc tttgtgtgta gcgtgtttt 46560
tcctcacact gacagctgta ggctgggggtt ctgagtgagc cccacagggc agaggcagaa 46620
aatgagtctc agagaggggtg agcgagctgc ttggggcccc acagcaggag atggagcagg 46680
actgcagcct agcctctgcc cccagcacct gcgcaagaag ctgctctgct ctggactgtg 46740
ttaggctgcg agggctggag agaatgaga gttgtgctt agagaggggg cgcagggtccc 46800
catggctttt cctcttatga tgaggtgat gggtaaggg aggggccatg ctgcagggg 46860
ccagtgaccg agggccgccg ttggaactga tggccttcat cccgagccca gccaggtgg 46920
gagcagggtt ttccgagggc ttgtctggg tcggcctgct tccagggact ctgctgcagc 46980
tcccaccct gtccaaagca tggaatcccc caggctccct ggagctctg tcaacctctg 47040
tcctccaag ctgagtgtgg ggcaagtct ggaggtcagc actgctcagg ggggccacg 47100
ggctgcttc aggggccaac cgcctgacct tggaggacgc caacatcgtg cagcctctgg 47160
gcctgacct ccttggaag catctctact ggatcgaccg ccagcagcag atgatcgagc 47220
gtgtggagaa gaccaccggg gacaagcgga ctgcatcca gggccgtgtc gccacctca 47280
ctggcatcca tgcagtggag gaagtcagcc tggaggagt ctgtacgtgg gggctggcag 47340
tgggtgggc aggggtggcct ctaaacccga cccctggagg aggctggagg ccagtgaag 47400
atcctgtgtg gcctcagcca ggcggtgtc tctgccagat gccaaactgt gcccgctggg 47460
gttcagcgac atgtccgaat gtcccaggc ctctgaggtt gttttttt gccgcagaac 47520
aatcaccac gaacagcgtt ttaagacaac accaactctt tttttttt ttttttga 47580
gtcaggatct tgctctgtt cccaggctgg ggtgccctgg tgcaaacaca gttactgca 47640
gcctcgacct ctgggcttaa ttaagtgaac acctgcctc agcctccag gtagctggga 47700
ctacaggtgg gcaccaccac acctggctaa tttttttt tagagacggg gtttcccat 47760

gttgcccagg ctggtctgca actcctgggc acaagctatc tgcctgctgt ggcctcccaa 47820
agtgctagga ttataggtgt gagccactgg cctgacaaca cccacggatt gtctctcagt 47880
tctgtaaggc aaagtccagg cacagcgtgg ctacactggg ttctctgctc agggctctac 47940
ggggccagaa tcaaggtgtc aggaacgtg ggcctcagc ggaggctctg tggagaaatt 48000
agcttccttg ctactcagc aggtagcagt tgtgggatcg aggttctgtt ttctctctgg 48060
ttattgctg gggaccactc tcagctccta gaggccaccc caggtccttg ccccggtggc 48120
ctctctgcct cagcagtggg ggctccctgc gtcagtcctt cccgcacctt gagtctctct 48180
gatttgcttc taaagggccc tgtgattcgg ctacgccacc tttagattag gttagcctcc 48240
cctttgatag actccaagtc ggctgattaa taaccttact cacatctgca gaatcccttc 48300
tgccacataa ggtcatgacg ccgtgctggg gactgggggtg ggaaattacg gggtcattta 48360
ggattctgcc tgccactgcc ttgtgtgtc ccagggtctg ggggaggggc ctccacagct 48420
gggaccacag tccttcctcc cctccatggt aacctatga ggattactg agaccagcct 48480
gggcaacatg gtgagaaccc atccctacaa aaaatacaaa caaaaaggga ccaggctggg 48540
cttggtggct catgcctata atcccagcac ttggggagac caagggtggc tgatcacttg 48600
agggtgggag ttcgagacca gcctgcccac catagtgaac tcccgtctct actaaaaata 48660
caaaaattag ctgggtgtgg tggcaggcgc ctgtattccc agctactggg gaggtcagg 48720
tgggagaatt actgaacct gggaggcgga agttgcagtg agccaaaatt acgccactgc 48780
actccagcct aggcaataga gtgagactcc gtctcaaaaa aaaaaaggg ccagggtgg 48840
tagtgacaaa gagaccctat cccaaaaaaa ccgaacactg aatccttgag actgagtaag 48900
gacactgtga aattttctg ggtggggcag ggaacagagc gtctctgtc atttctcca 48960
cctgggtgtg gtcagctctc cctccaagct gcctcctctt ctctcattg tccgggtgtt 49020
ggacacattt ggtaactgg atagaataac gcgagttccc agggacttgg tccatttgct 49080
attttattt attttattt attttattt attttattt ttattattt attttattt 49140

tgagatggag ttctgtttt gtcgccagg ctggagtga gtggcgcgat ctcggttcac 49200
tgcaacctct gcctcccagg ttcaagtga tctctacct cagccttcca agtaactggg 49260
attacaggca cccaccacca taccaggcta attttttgt attttagta gagacgggtt 49320
ttcgccattt tgcccaggct ggttttcaac tctagcctc aggtgatcca cgcacctgg 49380
cctcccaaag tgctgggatt acaggcatga gccaccacgc ctggcacat ttgctatttt 49440
aattcccatg tgtattagt tcccacggct gctgiaacaa atgaccacaa actggatggc 49500
ttaagcaac agaaatggat tcccccaatg tgctggagac cagaagcctg cgaccaaact 49560
gttgggaggg ctgtgcttcc tctgggggct ccaggaggga tctatttgtt ggcccttcca 49620
gtgctgtggg tgccagcgtt ccacacttgt ggatgcgcg cctcaacctc tgcccatctt 49680
catgtgtcca tctctttgt gtctgcgtct ttaccttctc ttctgtctg tgtgcctct 49740
tataaggacg ttgtcattg ggttagggc ccacccaaat catccgagat gacctcgtct 49800
tgagatcctt aacctgcaaa gaccctttt ccaaaaaaag gttatgtca cagattctag 49860
gccttaagac atgggtgtat ctttctgggg ggcactatcc aacccttat acaatgaaag 49920
acgggaagag ggccagggtgt ggtagttcac gcctgtaatc tcagcacttt aggaagctga 49980
agcgggagga tcaattgagc ccaggagttt acaagtagct aggaacatg atgagacccc 50040
atttctacaa aaagtaaaaa aaaaaaaaaa aaaaaaaag ccagggtgtg tggtcacac 50100
ctgtaatccc agcactttgg gaggtgagg caggcagatc acgaggtcag gagattgaga 50160
ccatcctggc taacacggtg aaacccgctc tctactaaaa atacaaaaaa ttatggccgg 50220
gcgcagtggc tccgcctgt aatcccagca cttggggagg ccgagggtggg tgaattacaa 50280
ggtcaagaga tcgagacat cttggctaac acggtgaaac cccatcaaga tcacaaggctc 50340
aagagatgga gaccatcctg gctaacacgg tgaaccccg tcttactaa aaatacaaaa 50400
aattagccgg gcatggtagc gggcgctgt agtcccagct gctcgggagg ctgaggcagg 50460
agaatggcgt gaacccggga ggcgagctt gcggtgagcc, gagatcgctc catgccattg 50520

atgggacgtg ctgacaggtc ctctgccggg ttctgcctt gctatgcgca cgctggtcac 54720
cacagaggcc tggcccttct tctgtagcag tcccacaccc gcaacagggtg tggctgctga 54780
ccacctgctt tctgcccctc tggctctgag gagggcgagc tgggcactca ggcgtggctg 54840
agcagatgtg tgttgccggg aggaggaagg actgctccag tcagggtga attcccacc 54900
cggagcattt ctgctgtatt tgggtagcg cctgctgctt aaagctctga ttcccagttg 54960
gcacccttcc cttctgcat tgaaaaacat acggatgcat gtctcttg agtgaatgtg 55020
tatttccca gcctctctc tgggttgggg ctggagggtg agcggcacac aggagccgca 55080
gcgatggagg atgtgcgggt gcagcacccc gtacagcagg gatgccaaac ccgcgctgag 55140
tccctctcaa cttctgctt gaagcccagt cacgccattg cctgggttt gctgggcggg 55200
gctgcatgtg atgttctct ctgtccctcc ccagagccg ccacctgct ccccgacca 55260
gtttgatgt gccacagggg agatcgactg tatccccggg gcctggcgct gtgacggctt 55320
tcccgagtgc gatgaccaga gcgacgagga gggctgcccc gtgtgctccg ccgcccagtt 55380
cccctgcgcg cggggtcagt gtgtggacct gcgcctgcgc tgcgacggcg aggagactg 55440
tcaggaccgc tcagacgagg tggactgtga cgtgaggcc ctccccgtca aggtctgccc 55500
aagaccctgg ccctgcccctc cgggatacga gcttggggct gcctccggcc tcacaggagt 55560
aggggctctg aaaaccttg ctgacaggga gattgccaag tctgtcttt aggcccaaca 55620
aggaaaactc tgcagttcca cccatcctgt cccaccaggt agtgtggctt gaaggcagac 55680
tgtgagggtc tatctcacct tctgcatta ggtcaggagt ttacagaaa cctgaggcac 55740
attcaggggt gggctgcaga ggtccatggc tcacaccctg gaaaatccgc cccaaaaga 55800
cagtgtgtc tccactgacc agtctgtggg atagtgtta agcctgagt gtttctatca 55860
acatgtagaa tcaggaggta taaagagatt tgctcaggca tctggggccc tctctgacca 55920
gcaggatctt ccttagatc ttgacagtga aacacatctc ttctgtccc cctgtgagt 55980
tttttcatt cattcattca ttattcatt cattcattca ttgagacag agtcttgctc 56040

agcagagggg acatttttg actgtcccc tcctgagctt cccagcagct ttctccaag 53340
ttacagccca aaagctcagg tggatttga acccaacggg gtctgtgcac ctcccactga 53400
tgcccgaaact gccctggcca agaaacgggg ccgtcagaac gctgcactaa ctgcagcctt 53460
gggcctccat gccagaggcc atgcccttc atccaccacc cctggcctg ggccctggcc 53520
ctcctggctc gggaactcca ggcccttc tcacggatcg agagacgtgt atttaccga 53580
caggtgcttg tcattctctt gtggcctctt ctccaggag atcacagaag gacaggcct 53640
cactgaggtc tcggacatgg acccttgat agtggcagga gccaggctgg gcaagaggcg 53700
gccacagtca cctcagcagt gccatcca ccgccattca gcccttcct gagccgggcg 53760
cgccctggc tctggcccca gtgtccagt tacagctcac aggagcttg ggtgccagc 53820
ggctgcttct gattgagagt cgaggcggg ggcttggga ggctgagagg ctgctcggtt 53880
tcacaactgc tgaggagac ttgggtcca tctcaggtt gcccatgtc gccctcaacc 53940
tccagccacc ggtcctccgt gtcccccattg gccaggcac gcttgacac atctgtcgtt 54000
ggctcctctc agccgtcgtg ggctgacctt ggcacgtcct cctgtggctg agcccagtgg 54060
ggacagctgc ttcctttat taccctagaa ctctcgtt tgatcaggcc ccctccccta 54120
tgccacacag tccctgtcac tcgggtgagc ccagtagtca tggggaaggc ctgcgggttc 54180
caaacatcca aaggcttgcg tgcagcatga cagcttgaac ccgatgttt ttacctgat 54240
cagattcag ctggcgggg gcttctca gcttcagt aggcctgggc cgatttcca 54300
gcatccctc ctgaggccag cctctgttc ctgtattt ctgcacaaag tgggaggag 54360
gagtcttagg aaatggggg ccacctgaa acctaggcct cctctggctt ctctgtcca 54420
gtgccccac gcttgtgtc tgtgtccca gccatggga ctgtgtatt ccctgagtgc 54480
tgccgcatgc ccagcccgca ctgaggacgt ggagccccga ggggcaggat ggcctccatg 54540
gtcacacgta ggaagtggc tccaccttc gatgacctc tccccctc ccttcagcg 54600
cctccccgg ggggtgtatc agccctcctg cctgtgctt gtccgtctt ctgcaggcg 54660

cactccagcc tgggtgacag agtgagactc cgtctcaaaa aaaaaaaaaa aaagaaaatt 50580
agccaggcac agtggcaggt gcctattgtc ccagctactt gggaggctaa ggcaggagaa 50640
tggcatgaac cggggaggtg gagtttgag tgagccgaga tcatgccact gcgctccagc 50700
ctgggcgata gagcaagact ctgtctcaaa aaaaaagcc aggcattggtg gtgcatgcct 50760
gtagtcccag ctactcaaga ggctgaggca ggagggtgtg tcgaccacg gagatcaagg 50820
ctacagtgag ccatgatgc accactgccc tccagcctgg gtgacagagt gtgaccctgt 50880
ctcaaagtaa gtaaatagga ggagagacaa gtgggcagti cagactgatg gtatgggcac 50940
agtagagact ggtgcagaca ggctggcctg tgatgtcaag caacttctgt aactgtttcc 51000
ggcatccatt tgtgtgtcaa ttccgtgtc agtaggaaga ctcttaggc tgccaagagg 51060
aataagtggg aggatcctcc cagagaggcc gggcctgcag gagggccagt tctcatgagt 51120
tcttatttg cccctaccct ccaggctgtg gttctgaggt gggagacaga gcctgacctc 51180
tgtttgtctt gttttgtctt tgcagcagcc caccatgtg cccgtgaca tggtggctgc 51240
tcccacatct gtattgcaa ggtgatggg acaccacggt gctcatgcc agtccacctc 51300
gtgtcctgc agaacctgct gacctgtgga ggtaggtgtg acctaggtgc tccttgggg 51360
tgatggacag gtacctgatt ctctgcctgc taggtgtctg cctggcatcc tttaaaatc 51420
acagtccctg tggcatccag ttccaaagc tgattgtgtc ttccttgcc ctctttctt 51480
ttctactatg tgcattcggg gctatgaatt ttctctaag tactgcgtt cctgcatctc 51540
acaaatttg ttacatttc atttcaggt agttgaata ttttacact tctcctgaga 51600
tgacatcttt ggctcatgtg ttatttagaa gtgttgcta gttttaaag agttggggct 51660
ttccagctg tctctctgca actgatttct aatttaattc tactgtagtc tgagagcta 51720
tttatatga ttctgttat tttaaagtgt ttgggtgtgg tgttttgtt gttattgtt 51780
ttgtgtctt ttgtttgtt ttgcttctt ttgtttgtt ttgagacagt gtctgtctt 51840
gtcactcagg ctggagtga atggcgcgat ctacgtcac cgcaacctct gcctcccggg 51900

tgtcaccag gctggagtgc cctggtgtaa tctcggctca ctgcaacctc tgcctccagg 56100
gttcaatcga ttctctgcc tcagcctccc gagtagctgg gatgacaggt gcgcaccacc 56160
atgcctggct aatttttga ttttagtag agacagggtt tcacatgtt ggccaggctg 56220
gtctcgaact cctgacctca ggtgatccgc ccgcctcagc ctcccaaagt gctgggatta 56280
caggcatgag ccaccgcgcc cggcctgagt ttctttta tgaaggacct gcttggttg 56340
ttgcctgcca catgttgta gcacatggg ccaggactg ctgaggagct gttgatgcc 56400
tcgtctccc agagccaccg gctctgttag ataattcaca tgcagtctgg cactgtcct 56460
acgtcctcat tcacaaagag cagacattc gtagaagatg agggcctggg agtaacctcc 56520
ctgcatgtt ttctataaag gcatagtgt taagtcttc cagctcattg accattggag 56580
aattttatg aggctgtaga ctaggggctg gtaaactaag ggcccagggg ccaatccag 56640
cctgccacct actttgtaa ataaagttt ctggtgcac agccatgcc attcattcat 56700
ttgcacaatg tctgtggctg cttcatgcc aaaagcagga gaactgagt gttatgctg 56760
agacctacgg cttcaaagc ccagacctc acgtctggc cttgacagac agagcttccc 56820
cagccctgct gcgcctctg gccagcatg tgctgtgtg gtgattcag cttgcaggag 56880
ccgtggtag gaattgtccc tgtgtggc cattttgcat tgctatgaag gagcacctga 56940
ggccgggtag attatgaagg aaagaggtct gctggctca tggttctgta ggcagacca 57000
gtatggcacc cgcctctgct cagcttctag tgaggctca ggaagcttg actcatggtg 57060
gaagtcgaag cgggagcagg tgcacacat ggtgagagag ggagcaacgg agagagagag 57120
agagagagag agagcgcctc tccctcttgc cctcacctg agaggagatg ccaggctcct 57180
ttaagtaacc agtcccatg tgaactcaca gtgagagccc atttgctact gcggagaggg 57240
caccaggcat ctgtcccat gacccaaaca ctgccacca ggccctacct ccaaccttg 57300
ggtcatatt tattctgtc tatgctatgc tatgctatgc catgcatgc catgcatgc 57360
tattctatt ctattattg agacagaatc tcgtctgtt gccaggctg gactgcagt 57420

gcatgatctt ggctcactgc aacctccacc tcccagggtc aagcgattct cctgcctcag 57480
cctccccgagt agctgggatt acaggcacac accaccacac ccgggtaatt tttgtatttt 57540
caatagagat ggggtttcac catgttgcc aggctggtct caaactcctg gcctcaagtg 57600
atccacttac ctgcgcctcc caaagtgcc tgattacaga tgtgagtcac tgcgcccagt 57660
gagggtcaca ttccgttga gatttgagg ggcagacgtt ggagccatct gagccccctc 57720
gtcccgtct agcttctct cccgtgtgcc ccgcgggtgt ggtggcaggc cttacgccg 57780
gttctggctg cacgctctgt tccagaagct ttctccctg ctggttacc agaaatcat 57840
cccatccatt acaaggacag ggtcccccta tctccattc ccagggcagg acaccggggg 57900
cagggcaggt ggggaactga gcaagttctc tgggggcagg cgtggctatg gctccctctg 57960
ggtgggcgtc tggggagggg tggaggcagc cgtcagcgcc ctggttctc cttcctcct 58020
ggccagagac tgtggccttg tgctgtccc gtgtgggtg cctgcacctc cagtgggttg 58080
tgctccctcc cctccctcc cctcaagctc tgctgagcac cactgcctc cacagcccc 58140
actctcggga ggcgaggctc ctctggcca ttctgtcct tggcacccac cccccacca 58200
acctggtaga gccttgggag ggtctgtta ctcttgcct ggcgtagacc tccccacagt 58260
aggcacctga cacatactc ctggggggca ggcaggaggt gcgttaggt ctacgccctg 58320
gcagtcctc ccttgcgtg cataggcctc gccacagggt catcagggt gggtagagac 58380
tgtactagac cactccccgc tggctctaga aagggtcca tctgtctgt cctgtttgg 58440
agtccagacc ttggttctg tgccctgcat ggtgggtg ggggcacct ccagcctctc 58500
tgagtcatg gcctctctt gcagccatct gcctgccc aaagttccg tgtgcgagcg 58560
gccagtgtgt cctcatcaa cagcagtgcg actcctccc cgactgtatc gacggctccg 58620
acgagctcat gtgtgtgag ccagctctg gcacggggaa gggcggtccg ggctgggttc 58680
ccccaggaac gtggagtta ggggaggaga cgtgccttc cagcggggt gggggctgtg 58740
tgggagactc aggcggctg gaggtcctt gcgggaggca gggaagcct tcccaggga 58800

ggcttatggg tggcgtgaat tagtcgggggt ctatcaggag gcagaaactc tatgagaatt 65760
tgaacagaga aagttccgtc tacaggctta ttaccaggga ctggaatagc agaaattgaa 65820
cagtgaatg tacagagaac tctaagaatg caggaatagg ccaggcatgg tggctcacac 65880
ctgtcatccc agcactttgg gagaccaagg cgggtggatc acctgagggtc aggagtgcga 65940
gaccagcctg gccaacatag tgaaacccca tctctactaa aaatacaaaa aaattagctg 66000
ggtgtggtgg cgcgtgcctg taatcccagc ttctcgggag tctgaggctg gagaatcact 66060
tgaacctggg aggcagaggt ttagtgagc cgagatcatg ccattgtact ccagcctggg 66120
caacaagagc gagactcagt caaaacaaca acaacgcagg aatagcagat gagccgaggt 66180
ggggcctccc cagccccac cccccaccc gcaccctggg ccgagatcca gtcctctttg 66240
aatagggcct gggcgtggtt cacgggacat ctgagacatt gccgaggcgc tgcactggtg 66300
gatctgcca gaagtctgcc cagtgcagat ttgggcagaa tctcaaactg ccttgggatg 66360
taggagagaa accaggcctg gtcaagtcca tgggaagagg tggaacaga ccccataggc 66420
tggggcttgg gcagctgtag gaagccctct ctgtgcctc cctgcctgct ctctgctttg 66480
aagcatcttc ccagtgccc ccagtctcat gccctctcaa cgttggggtc aaatcctgag 66540
gaataccag actggctctc tgggcaaag aggaccctct ccagaaagag cagggccag 66600
tgcggcttc taaagggcag gggaagggcc tggccactcc ccagaggcta ctcaccagcc 66660
atcaggatag cccaggaag caggcctct cgagccatt ttattactt atttattat 66720
tttattaat tttaaattta tttttgaga cagagtctca ctctgtgcc caggctggag 66780
tgcagtgtg cgatctcaac ccactgcagc ctctgcctcc aggggtcaag ggattctccc 66840
acctcagcct cccaagtagc tgggattaca ggtgcccgc accacaccg gctaattttc 66900
atattttag tagagacgag gttcacat gttggccagg ctggtctga actcctgacc 66960
tcaagtgatc cgccgcctc ggcctccaa agtgctaggt caagccatt ttaaagttga 67020
agaaactgag gctgaggtaa attccctccc caggatcct gctgcagcca gaaggtgta 67080

ggctcccagc tgccgtgtgc accctgccta gagctctacc gtaacccatc tccgggagga 64380
gggtgctattg tttcctcat ttgcaacaa ggaggctgaa gaactgagca tgaaccactg 64440
gcctgggtcg ttcggttggt aggcagtggg gccaggccat ccaactcaca accaccttct 64500
actctgcttc ccccgacccc tgaagtttgt tctgtttga ggacacagcc gtcacattct 64560
tggtggctga acagcactcc ttgtcagggtg ttgctgggcc ccactggag ggcatcatgg 64620
tcctctctcc tgctcgggtt gaacctggc tgttcaacc actcctgcca agtggccctc 64680
tgaaagggac agtccatctt ttctcagcag agggccacac tggcaaacg gtcctggca 64740
ccctttctct ccacctgtct aatatagagt aaaaatgta tcatgttaag atcttcattt 64800
atatttattt tatcatgaat gatgtaagca tcattttgtg tgttaagaa ccttgggcc 64860
cagcgtgatg gcttgcagct gtaatctcag cacttagga ggctgagatg agcggatcac 64920
ttgaggccgg gagtttgaga ccagcctggc caacatggag aaacccgctc tctagtaaaa 64980
atttaaaaat tagccgggta tggatgccc agctacttgg gagtctgaag catgagaatt 65040
gcttgaacat gggaggcgga ggttgcagtg agccgagatc gcgccattgc actccagcct 65100
gggcgacaga gcgagactct gtctcacaaa aaaaaaaaaa aaagaaaaga aaagaaatta 65160
tcaatctcct ctttatggc atatatatat atatatatat atatatatat ttattccct 65220
ttcttggtta tgtcataaa ggcctcccct gctctgatca taaaaacaa cttatttca 65280
cactctctct cttttttt tgagacagag tttgctcct gtgcccagg ctggagtga 65340
gtggcgcaat ctcagctcac tgaacctcc gcctcccggg ttggagtgat tctctgcct 65400
taccttcccg agtagctggg attataggca tgcaccacca tgcctggcta atttgtact 65460
tttagtagag acggggggtt ctccatgttg gtcaggctgg tctgaactc gcgacctcag 65520
gtgatccacc cacctcggcc tcccaaagtg ctgggattac agacgtgagc caccatgccc 65580
agccacact ctctttcta acgtcctcct ccttcgttt tacgttcaca tctttaattc 65640
ttctgggatg taattagatt tgatgagcaa ggtgggcac cagctgttt cttggctgat 65700

cggggcccggg gaggggcccggg gcgggatggg gctgtgggcc cctcccaccg tcagtgtctgg 61620
ccaccggagg ctccccgggt tcctgggggc tgtgccaccg cctctgaggc atgcttgctt 61680
tcttcccttt tcaaaccctt ctgttctt cttaatgac attgttgatt gtggataatc 61740
tgaaaactac acaaaaatat aaagagccaa aatctcacc aaatccacct cctagagtgg 61800
ctgttgggct ccgtcagcat ccaggcggcc gtctgtgttc cgcacggccc agcccatcga 61860
tagccgcctg caccaggcct gtctgccctc tgtgagcctc cccacagggt tccctccaca 61920
aacaccctgt tctccaccc agggctggct gcttctgga aaacagtgg atggtttgt 61980
gcatgacaga caaacacagg gtgatttcg tggctaaaat actccctgga gcttttgga 62040
gggtgagggg ctggctccag ctgagccacg ccttgagtga aatgactgtg aggagaataa 62100
actgccctg cctccagga tcactggggc tggctgggga gaacccccgt ttctgggagc 62160
acagtcccag gatgccaagg cgagcttggt gccgagatgt gaactcctga gtgtaaacag 62220
cgggggctga ctgacatgc ttgtatgct ttctattgt tctgcagct gtatgccct 62280
aaggtgagtc cagccccctt ctgttctc tggggcctc cagtgagcc ccacctgct 62340
ggggctggtt cctctgccc ttctgggtat cctcacatc tggggtctg tcttctgtt 62400
ttattttct tttttttg agacggagt tcactttgt tgcccaggct tcagtgaat 62460
ggtgtgatct ctaggctcac cgcaacctct gcctcccagg ttcaagcagt tctctgcct 62520
cagcctcct agtagctggg attacaggca tgtgccacca cgccagcta atttgtatt 62580
tttagtagag atggggtttc tccatgttg tcaggctgat ctgaactcc ctacctcagg 62640
tgatccgcc accctggcct ccaaagtgc tgggattaca ggcgtgagcc accgcacctg 62700
gccttttct ttttttct tttttttt ctgagacagg gtctcgtct gtcaccagg 62760
ctggagtga atggtgtcat catggctaac tgcagcctct acctctagg ctcaagcaat 62820
cctccatct cagcccctaa gtagctagga ctgcagcat gcatcccat gccagctaa 62880
tattacatt tttgtagag atgaagttc actatattgc ccaggctggt ctccaactcc 62940

aaacaggact tcacccgggt ctgtctggcg tgaaggcag tgttcttga ccaccctagg 67140
gggcctgaga gaactgagtc cctcgggcat aactgacagt tctgttcca ttattccgca 67200
ggggctcgga tctggctgta tgctttccag gatggccttg gagaccaca taagccctac 67260
accctttggg aagctgcatg ttgggttggg gtgccgtcag tggcacttgt ggaaggtgca 67320
gacctgtgtg ggtgtgtggg cccagggccc ctgtccctt cctcccttg tagggctggt 67380
tgtgtgctgc ctggacctgg ggggcacgtt cacgtggtga attgtctat ttactatccc 67440
cgctttgggg ctggtgccag cacaggccct tgtgaagggg gtgccttct ctggagtggg 67500
actgtggccc ctccctcagc gtggtgactt ctgtgtcagg gcttcagcag ggacgcagag 67560
cccctgagtg ttggaacaa gggcgtcatt gcaggagtta gactgtgtgt gatggaggga 67620
ggagggggcag gaggaaaggc cagaaggaga gttcctggga aggtccctga ggagcctggt 67680
gaggtgctaa ctggttgga ggacactcag ggctgtggg gacatctct actgtgggg 67740
gccagccaca aagggaactg gccgaagtc tgtccccgc ttacagccc agcatctggt 67800
cacaaggcag gtacttgga gggcgcgggc acctgggcca aaagtgcctg gggtccctt 67860
gcctttact gagatgacct tcggggcagg tggctgctgc ctccctcct gtcccaggt 67920
tttgccaact ggccagagga aggggtcctg ggaagcaggg gggccagaag cctctctgc 67980
aaggaaagcc cgaggggtgt gggaggaagg aaggaatgcc caggctggcg aggctctaag 68040
tcacctggc ttggctctcc tcagatctg aaccgccc cctcccggc cacggacccc 68100
tccctgtaca acatggacat gttctactt tcaaacattc cggccactgc gagaccgtac 68160
aggtaggaca tccctgcag cctccatgg ccattgggt cccgccagcc cgtggtggag 68220
gggcctaate ccatgccac tgatgagggg aggtattctg ggtgctagt ggcaggtgcc 68280
ggggccagcc ctgcctcct ctgctctgcc aaccacata ggctgcctcc ccagacaagc 68340
tcagcgggca ctgcatgttg ggttcagaaa tcagcagaac tccagttct gagctgctt 68400
tcaagttgct cctatggggg ttacttttaa gctgggaaat ggcgtggcg tcgaggggcc 68460

ctaattaaag agacggggca gtggacaggc attttcagtt gactgccag ggagtgttct 180
gccaacagg gaggatatgc gtacagaatc atactgatc agcatgagtc caattcagac 240
cgtacatcag tggagatatg ggtcccccga tgactccgtg gaacactgat gttgtgaca 300
ggggagtaca gcaccagcca tcagcaggcc agtaaactat accggcctgc gaaattggac 360
tcagaccgg atccaccctg accgacgtcc caagccccc cccccaccc cccaccatgg 420
gccgagatcc agtctctttt gaatagggcc tggccgtggt tcacgggaca tctgagacat 480
tgccgaggcg ctgcattggt ggatcttgcc agaagtttg ccagtcaga ttgggcaga 540
atctcaaact gccttgggat gtaggagaga aaccaggcct ggtaagtgc atgggaagag 600
gtggaaacag accccatagg ctggggcttg ggcagctgta ggaagccctc tctgtgcct 660
ccctgectgc tctctgttt gaagcatctt cccagtgcc cccagtctca tgcctctca 720
acgttggggt caaatcctga ggaataccca gactggctct ctgggcaaa gaggaccctc 780
tccagaaaga gcagggccca gtgcggcttc ctaaaggga ggggaagggc ctggccactc 840
cccagaggct actcaccagc catcaggata gcccaggaa gcaggccttc tcagcccat 900
tttattact tatttatta tttatttaa tttaaattt atttttgag acagagtctc 960
actctgttc ccaggctgga gtgcagtgg gcgatctcaa cccactgcag cctctgcctc 1020
cagggttcaa gggattctcc cacctcagcc tccaagtag ctgggattac aggtgcccgc 1080
caccacacc ggctaatttt catatttta gtagagatga ggttcacca tgttgccag 1140
gtggtctcg aactcctgac ctcaagtgat ccgccgcct cggcctccca aagtgctagg 1200
tcaagcccat ttaaagttg aagaaactga ggctgaggta aattccctcc ccagggatcc 1260
tgctgcagcc agaagggtgt aaaacaggac ttcaccggg tctgtctggc gtgaaaggca 1320
gtgttctgt accaccctag ggggcctgag agaactgagt ccctcgggca taactgacag 1380
ttctgtccc attattccgc aggggctcgg atctggctgt atgcttcca ggatggcctt 1440
ggagaccac ataagcccta cacccttgg gaagctgcat gttgggttgg ggtgccgtca 1500

ctgtcatcca ggctggagtg cagtgggtaca atctcagctc actgcaagct ccgactccca 71280
ggttcaagtg agtctcctgc ctcagcctcc cgagtagctg ggactacagg tgcgcgccac 71340
cacacccgcc cagctaattt ttgtattttt agtagagatg gggtttcacc atgttgacca 71400
ggatgatctc gatctcttga cctcgtgatc cgccacctc ggctcccaa agtgctggga 71460
ttatagcat gagccactgt acccagctga ctcttagtca ctttaagaa ggggactgtg 71520
ccttcatttt tcaactgggcc ctgcagaata tatgcctggg ctctgggctc ttctgaacct 71580
gtgttggtt ccatctgacc tctctgtgcc agcccaagge tgctgcttt cctgaggga 71640
aggagcccca tgactgcgtg ttgactcgct ggatggggct gctgagcca ctctgccaca 71700
ccacgtgccc ctggcaggga gggaatccct ggtcctcac aggaacagtc agcaagccac 71760
acctgacgcc tgctgtgggc ccatccctgc ggtgctggag aagacagaca aggctggtc 71820
actgcctctg caggggtccc agtccgtgga aggagacagt aatctaggca tttcgggtgg 71880
ggaagctgag ctgttctcgt gtcctgaagg ccaggcggga acagccgtct tcagaggga 71940
gggagaaaat gcacatgca tcagtggaga agggcctgac ttcctcagc atggtggagg 72000
gaggtcagaa aacagtcaag cttagtatt ctatagtgc acctaaata 72049

<210> 10

<211> 8705

<212> DNA

<213> Homo sapiens

<400> 10

ggactcaggg gcagcaggga ggtacacca tggtagtgg gcggaccata gggggtaatg 60
agagggtgaa tcgatggaac ctgggggaca caatcgaagt ggtccagag tcgggctgta 120

gaaccacccc catgattcaa tcattctcca ctgggtccct cccacagcac gtgggaatta 69900
tgggagtaca attcaagatg agatttgggt ggggacacag ccaaacccta tcggttgcca 69960
acatttacag taacagtgtt aggtgaacag ttgtccagtc tcctgtttg tcggacactg 70020
tttctagcac ctccaggca gaatctcatg tctccttcac ttctgaaatg ggtactattt 70080
catccccact ttatcaatg agaaactaaa gctcgaagag gtcaagtaag ttctggcca 70140
aggtcagcta gcaggctcta gaggcctcgt tctccttaga ggcagccttg ccagggccca 70200
ggcttggcag gctgcagggc aggtgcgggc atgcccattg tagaggtggg accattgagg 70260
ctcagagagg gtaagtgatg agccctggcg acacagcggg glgggtccag agtccggcct 70320
gcatcttctg gagctggcca gtggacaggc ctttccggtt cacagccccg gggctgctgt 70380
gcccaccagg gcggatgtgc ctaccgaatc cactcctct gtgtgtgtcc ctttcaggcc 70440
ctacatcatt cgaggaatgg cgccccgac gacgcctgc agcaccgacg tgtgtgacag 70500
cgactacagc gccagccgct ggaaggccag caagtactac ctggatttga actcggactc 70560
agaccctat ccacccccac ccacgcccc cagccagtac ctglcggcgg aggacagctg 70620
cccgcctcg ccgcccaccg agaggagcta ctccatctc ttccgcccc ctccgtcccc 70680
ctgcacggac tcatcctgac ctggccggg cactctggc ttctctgtgc ccctgtaa 70740
agttaaata atgaacaaag aaaaaatat atttatgat taaaaata aatataattg 70800
ggattttaa aacatgagaa atgtgaactg tgatgggtg ggcagggtg ggagaactt 70860
gtacagtga gaaatattt taaactaat ttgtaaaac agaactgcca ttctttgtg 70920
ccctgtgtgc attgagttg tgtgtccccg tggagggaat gccgacccc ggaccacat 70980
gagagtcctc ctgcacccgg gcgtccctct gtccggctcc tgcagggaag ggctggggcc 71040
ttgggcagag gtggatatct cccctgggat gcatccctga gctgcaggcc gggccggctt 71100
tatgtcgtg tggcctgtgc cgtcagaaag ggcctgggc tcatcacgc tgttctgtt 71160
cgtctctc agattcttag tctttttt tttttttt tttagagcg gactcttct 71220

gcggccagga ggacagactg tgagctgtgg gctcggcggc tacagagtct gcctcagtgg 58860
gcggggctga tgggtgccag gtgcctgcag cacgcaccca cccacgggac ctgtctgagc 58920
agcgtctgtc aggcagcaag attacccgag ggctgcagtg gtcctgttcc ctggcagctt 58980
actgtctggc tgaggaggag tgatgttcac atatgcacac atgtcatgtg cacacacatg 59040
tacatgacaa catcccatat gtcctcaaa tagcatgacc tgtacagtca cggatatagg 59100
gcctagggga taggaggcca agacagtcat ggaagacttt ccagaggcag tggctcctga 59160
aaggctgtct gattcaggca ggaaggaggc tgagttcaga taggaagtag caatgagtca 59220
ttgtgtctgg ggacatggcc actccttcgc tgcagaggga cctgggctga gagctcctct 59280
cttatggctg cagtcgggag agaagtctgt tggggggaga agggggcttc ctcaaggagc 59340
tcctgtgcc ctttggcacc ttcgtgccag gtcaggcttg aggcctgaag gcagtgggtg 59400
gggccaccaa gggtcgctc ctctgtggg caagtcca gtctgacggg cctgtgccgt 59460
gggccccagc tgtgggggag ctgttgatgc gcagccaggc ctgcccga gagcccgac 59520
gcttcattc cgtgacttc atcgacgcc tcaggatgc tgggccggc ctgtgggaga 59580
gtgaatgtgg cttttgcaa agttgagtct ggagcctgga aacttccta tgggcagcct 59640
tgatagtgga gtggcccaag gagccaccc agccgacct gccctcccg tggctgggtg 59700
gcggcaccag gggctgcctg gctttgctcg ttcaccaaca tcaccgggc tggccagggc 59760
gcgtcactt ctgccaccac cgagggccct gggcgaagga gtgaatacca ggctgccttg 59820
gcagggatgt gttgagggt gtggggagtc ggacagcggc gggggtcaga ggaggaggag 59880
ggtgcaccgt gcaggctgaa gggccacgt accctgaggt tggccaggct cccaggcct 59940
agcctccag ctccccact ttctccac cctccaccag tggcaaagcc agcccctca 60000
ggg'gcacgg tgtctgccc caaggaggc cattccgtt ggggttaatg ttggcacct 60060
ctttctgtt gtctctggca gaaatcacca agccgccctc agacgacagc ccggcccaca 60120
gcagtgcat cgggcccgtc attggcatca tctctctct ctctgcatg ggtggtgtct 60180

gtggcacttg tggaagggtgc agacctgtgt ggggtgtgtg gccagggcc cctggtccct 1560
tcctcccttt gtagggctgg ttgtgtgctg cctggacctg gggggcacgt tcacgtgggtg 1620
aatttgtcta ttactatcc ccgctttggg gctggtgcca gcacaggccc ttgtgaaggg 1680
ggcgccttgg tctggagtgg gactgtggcc cctccctcag cgtggtgact tctgtgtcag 1740
ggcttcagca gggacgcaga gccctgagt gttcggaaca agggcgatcat tgcaggagtt 1800
agactgtgtg tgatggaggg aggaggggca ggaggaaagg tcagaaggag agttcctggg 1860
aagggtccctg aggagcctgg tgaggtgcta actggtgtgg aggacactca gggcctgtgg 1920
ggacatctcc tactgtggg gccagccac aaagggaact ggccgaagtc ctgtccccgc 1980
cttcacagcc cagcatctgg tcacaaggca ggtacttga agggcgaggg cacctgggcc 2040
aaaagtcct gggttccctt tgcctttcac tgagatgacc ttggggcag gtggctgctg 2100
cctccctcc tgtccccagg tttgccaac tggccagagg aagggtcct gggaagcagg 2160
ggggccagaa gccctctctg caaggaaagc ccgaggggtg tgggaggaag gaaggaatgc 2220
ccaggctggc gaggtctaa gtcaccctgg ctggctctc ctcatcct gaaccgccg 2280
ccctccccgg ccacggacc ctcctgtac aacatggaca tgttctact tcaaacatt 2340
ccggccactg cgagaccgta caggtaggac atccctgca gccctccatg gccattgggt 2400
tcccggcagc ccgtgggtga ggggcctaata cccatgccca ctgatgagg gaggtattct 2460
gggtgctaata gggcaggtgc cgggcccagc cctgcctccc tctgtctgc caaccacact 2520
aggctgcctc ccagacaag ctacgcgggc actgcatgtt gggttcagaa atcagcagaa 2580
ctccacgttc tgagctgtc ttaagtgc tctatgggg gttacttta agctgggaaa 2640
tggtgtggc gtcgaggggc cgggggcttg ggctccagag tctgactgtg tgtttgagtc 2700
cggctgtgga aacctagcca ttgagatgcc cctcttggg ggctcgtcc tcttaggatg 2760
ggacaagtct gtgaaggctg ctgcagcacc caccgtagac cctaatacgt gtgacgtcac 2820
caggatggtc cgggctgtc acttgccaca gtggcctgtt tgagcccggg aagccaacgg 2880

tcggcctccc aaagtgcctgg gattataggg atgagccact gtaccagct gactcttagt 5700
cacttttaag aaggggactg tgccttcatt ttccactggg ccctgcagaa tatatgcctg 5760
ggctctgggc tcttctgaac ctgtgttggc ttccatctga cctctctgtg ccagcccaag 5820
gctgctgctc ttctgaggg caaggagccc catgactgcg tgttgactcg ctggatgggg 5880
ctgctgagcc cactctgcca caccacgtgc ccctggcagg gagggaatcc ctgggtcctc 5940
acaggaacag tcagcaagcc acacctgacg cctgctgtgg gccatccct gcggtgctgg 6000
agaagacaga caaggcctgg tcaactgcctc tgcagggctc ccagtccgtg gaaggagaca 6060
gtaatctagg catttcggg ggggaagctg agctgttctc gtgtcctgaa ggccaggcgg 6120
gaacagccgt cttcagaggg aaggagagaaa atgcacatcg catcagtga gaagggcctg 6180
acttcctca gcatggtgga gggaggtcag aaaacagtca agcttgtgc tgggtgacag 6240
tgcatttaat aatcaaaata taggctgggt acggtggctc atgcctgtaa tcccagcact 6300
ttgggaggct gaggcagggtg gatcactga ggccaggagt ttgagaccgg cctggccaac 6360
atggcaaac ctcaactact aaaatacaaa aactagccgg gcgtgggtgt gcacgcctgt 6420
aatcccagct actgggagg ctgaggcagg agaattgctt gaacctggga ggcggaggct 6480
gcagtgagcc gagattgtgc cactgcactc cagcctgggc aacagagcaa gactctgtct 6540
caaaaaaaaa aaaaaaaaaa gcaatacaaa atacaaatat cactttcact aaaagaaggg 6600
atggaagacc caaaacaac agaaaacaac aaaatggcag gagtaagtcc ccacttatca 6660
ataataacat tgactgtaaa taggctaagc tctgcaatca aaagagtggg ccaggagcgg 6720
tggctcacgc ctgtaattcc aacgctttgg gaggctgagg cggatggatc atttgatgc 6780
acgagtttta agaccagcct ggccaacaag gtgaaacccc atctgtacta aaaatacaaa 6840
aattagccag gcggtagtgg cacgcacctg taatcccagc tacttgtgag gctgaggcag 6900
gagaatcact ggaggctggg aagcggaggt tgctgtgagc caagatggag ccactgcact 6960
cccacctggg cgacagagtg agatcctgtc ttaagaaaaa aaagagtgga tgaatggatc 7020

gtaatgaacg tgaccaggag ctgcttactg aggacgcact ggatgatctc atcccttctt 660
ttctactgac tggtaacag acaccggcgt tcggtcgaag agtatctggt gtcataaaaa 720
ttgccgatgg gagtcgccgt cgtaaagctg ctgcacttac cgaaagtgat tatcgtgttc 780
tggttggcga gctggatgat gagcagatgg ctgcattatc cagattgggt aacgattatc 840
gccaacaag tgcttatgaa cgtggtcagc gttatgcaag ccgattgcag aatgaatttg 900
ctggaaatat ttctgcgtg gctgatgcgg aaaatatttc acgtaagatt attaccgct 960
gtatcaacac cgccaaattg cctaaatcag ttgttctct ttttctcac cccggtgaac 1020
tatctgcccg gtcaggtgat gcactcaaa aagcctttac agataaagag gaattactta 1080
agcagcaggc atctaacctt catgagcaga aaaaagctgg ggtgatattt gaagctgaag 1140
aagttatcac tctttaact tctgtgctta aaacgtatc tgcataaga actagttaa 1200
gctcagaca tcagtttct cctggagcga cagtattgta taaggcgat aaaatggtgc 1260
ttaacctgga caggtctcgt gtccaactg agtgtataga gaaaattgag gccattctta 1320
aggaactga aaagccagca ccctgatgcg accacgtttt agtctacgtt tatctgtctt 1380
tacttaatgt cctttgttac aggccagaaa gcataactgg cctgaatatt ctctctgggc 1440
ccactgttc acttgatcg tcggctgat aatcagactg ggaccacggt cccactcgta 1500
tcgtcggctt gattattagt ctgggaccac ggtccactc gtatcgtcg tctgattatt 1560
agtctgggac caggtccca ctgtatcgt cggctgata atcagactgg gaccacggtc 1620
ccactcgat cgtcggctc attattagtc tgggaccatg gtccactcg tatcgtcgtt 1680
ctgattatta gtctgggacc acggtccac tcgtatcgtc ggtctgatta ttagtctgga 1740
accacggtcc cactcgtatc gtcggctga ttattagtct gggaccacgg tccactcgt 1800
atcgtcggtc tgattattag tctgggacca cgatccact cgtgtgtcg gtctgattat 1860
cggctcggga ccacggtccc acttgattg tcgatcagac tatcagcgtg agactacgat 1920
tccatcaatg cctgtcaagg gcaagtattg acatgtcgtc gtaacctgta gaacggagta 1980

gggtgggtgct tcgtggcgtg gttctgaaac ttcgttgaa gtgtgtggac agtccttgc 3420
ctgttctctg tgggacccta ttagaaacg aggtctgagt tactgggggt catcactgtg 3480
ttctgatggc ccagctgtgt ggaggccgcg gtgcagcccc atccaaggag ccagggccct 3540
gggtctagcc gtgaccagaa tgcatgcccc ggaggtgttt ctcatctgc acctgtgtt 3600
cctggtgtgt caagtggctg tgaaactctg tgtagctct tgggttctt gaaagtgcc 3660
ccgggtctca ggcctcagaa ccagggtttc ccttcatctc ggtggcctgg gagcatctgg 3720
gcagttgagc aaagagggcg attcacttga aggatgtgtc tggccctgcc taggagcccc 3780
ccggcacggt gctggggcct gaagctgccc tcgggtgtg gagaggagg agcgatgaag 3840
tggcgtcag ctgggcagga agggtagacc cctgcaagg gggcatgctg gggacgtga 3900
gcagcatggc cagcagctgg gtctgcagcc tggtaaccgg cgggacttgt ggttggggct 3960
ggtttgtggc caggagaggg gctggcagga gacaaggggg actgtgaggc agctcccacc 4020
cagcagctga agcccaatgg cctggctgtg tggctctcag ctgcgtgcat aacctctcag 4080
tgcttcagtt ctctcatttg taaaatgagg aaacaaacag tgccagcctc ccagaggtgt 4140
catgaggatg aacgagtgac catgtagcat gggctgggtg cgtgtcacct aacatcacca 4200
gcctttgcaa ggagagccct gggggcctgg ctgagtattt cccttgcccg gccacccca 4260
ggcctagact tgtcctgct gcaggccctt gaccctgac ccattgcac ctgtctccac 4320
aggagccgag gaggtgctgc tgctggcccg gcggacggac ctacggagga tctcgtgga 4380
cacgccggac ttcaccgaca tcgtgctgca ggtggacgac atccggcacg ccattgccat 4440
cgactacgac ccgctagagg gctatgtcta ctggacagat gacgaggtgc gggccatccg 4500
cagggcgtag ctggacgggt ctggggcgca gacgtggtc aacaccgaga tcaacgacce 4560
cgatggcatc gcggtcgact ggggtggccc aaacctctac tggaccgaca cgggcacgga 4620
ccgcatcgag gtgacgcgc tcaacggcac ctcccgaag atcctggtgt cggaggacct 4680
ggacgagccc cgagccatcg cactgcaccc cgtgatgggg taagacgggc gggggctggg 4740

acctcgggtgt gcggttgat gcctgctgtg gattgctgct gtgtcctgct tatccacaac 2040
atittgcgca cggttatgtg gacaaaatac ctggttacct aggccgtgcc ggcacgttaa 2100
ccgggctgca tccgatgcaa gtgtgtcgtc gtcgacgagc tcgagagctc ggacatgagg 2160
ttgccccgta ttcagtgtcg ctgatttga ttgtctgaag ttgtttttac gttaagtga 2220
tgcagatcaa ttaatacgat acctgcgtca taattgatta ttgacgtgg ttgatggcc 2280
tccacgcacg ttgtgatatg tagatgataa tcattatcac ttacgggtc ctttccggtg 2340
atccgacagg ttacggggcg gcgacctgc gggtttcgc tatttatgaa aattttccgg 2400
ttaaggcgt ttccgttctt cttcgtcata acttaatgtt ttattttaa ataccctctg 2460
aaaagaaagg aaacgacagg tgctgaaagc gagcttttg gcctctgtcg ttcttttct 2520
ctgttttgt ccgtggaatg aacaatggaa gtccgagctc atcgctaata acttcgtata 2580
gcatacatta tacgaagta tattcgatgc ggccgcaagg ggttcgcgtc agcgggtgtt 2640
ggcgggtgtc ggggctggct taactatgcg gcatcagagc agattgtact gagagtgcac 2700
catatgcggt gtgaaatacc gcacagatgc gtaaggagaa aataccgcat caggcgccat 2760
tcgccattca ggctgcgcaa ctgttgggaa gggcgatcgg tgcgggcctc ttcgtatta 2820
cgccagctgg cgaaaggggg atgtgctgca aggcgattaa gttgggtaac gccagggttt 2880
tcccagtcac gacgttgtaa aacgacggcc agtgaattgt aatacgactc actatagggc 2940
gaattcgagc tcggtaccgg gggatcctct agagtcgacc tgcaggcatg caagcttctc 3000
ttgtgccggt tgtacgtgt caggtcacac tggtagtta ggcagggcac agatgccag 3060
agcagaggga actttccttg gggattcaac acgtgcaagt ctaggggct ggcaaatcct 3120
gccctcagct agagaggggg cttttattg agaccagaat cacctgagca tcctcctgtc 3180
cccagctgtg tccagcctgt ctgcaggac atcctgagag gaccaggctc tcccctcatc 3240
cacctgccta agtgccactc tgaaccctgt ccacctgtc cgtggagggg cgtgacctca 3300
agctgctcag ccagcagcag gcttggccct ggggggcagc agagaccag gtggctgtgg 3360

gcctggagcc agggccaggc caagcacagg cgagaggag attgacctgg acctgtcatt 4800
ctgggacact gtcttgcac agaaccgga ggagggttg ttaaacacc ggcagctggg 4860
ccccacccc agagcgtga ttcaggagct ccaggcgagg gctgaagact tgggtttcta 4920
acaagcacc cagtgttccg gtgtgtctgc tgggtccatg ctagaaagc cctggagacc 4980
tggaggagc ctttgttcc cctgggttca gtttctcat ctgtagaatg gaacgtcca 5040
tctgggtgat ttccaggatg acagtagtga cagtaaggc agcctctgtg aactgacca 5100
cagtacaggc caggcctctt ttttcttt tttttttg agatggagtc tcactctgtc 5160
gcccaggctg gagtgcagtg gtgtgatctc agctcactac aacctctgcc tcctgggctc 5220
aagtattct cctgcctcag cctcctgagt agctgggatt acaggtcct gccactgtgc 5280
ttggctaag tttgtattt tggtagagat ggggtttcac cgtcttgcc aggctggtcg 5340
caaactcctg acctcagtg atccacctgc ctacgctcc caaagtgtg ggattacagg 5400
catgagccac cagccccgt caggccaggc ctctttgaa cacttgac accatgggtc 5460
tttcatcca ggggggtagg tacagttga cagttgagga cactgaagc cagagaggct 5520
cagggacttg cccagggtca cacagcagga tgtggcaggt gtggggctgg gcctggcagc 5580
gtggctccag cttccagca tagaaatctg tgaagcaga tagttgtcg gtcggtaggg 5640
gagactttct gagaccgcc ccagcggctc agagggtagt agccaggggc ctctctgggg 5700
gtcataacc cagaacactg aatgggaaaa ccctgatgga ggaggcgag tggagctgtg 5760
ggtgccgatg ggaagtccca gaggagctgg gaggtcagta gcggtgtgc cctctgtgga 5820
gcacttagtg ggcaccaggt gtgtttccag gtcatggcc ctgggacctg aagctcagaa 5880
ggtgaagtaa cttgccagg gcacccgtcg ggcagcgcg ggagaggat ttgtgggctg 5940
tggagcctgt gctcgtggcc cagccctggg ggtgtgagt gtgctggcg gggagctttt 6000
cctgcaagtg gactggtgtc taggagccag catgtcaggc agcaggcagc gggagtgcag 6060
caggcagcg gagcacagca ggcagagggc ggggctcag cagccatccg tggaccctgg 6120

ggcacggagg catgtgggag agggctgctc catggcagt gctgaagggc tgggtgtgc 6180
cccaggagg gtgatgagg gtaagaagt gggccccag gggcttagc aagaggaggc 6240
ccaggaactg gttgccagct acagtgaagg gaacacggcc ctgaggtcag gagcttggtc 6300
aagtcactgt ctacatgggc ctgggtgcc tcatctgta aaaaggaagg gatggggaag 6360
ctgactcaa gggccctct agccctggtt tcatgagtct gaggatccca gggacatggg 6420
cttggcagtc tga~~ct~~ctgta ggtcgtgggg tccaggagg ggcaccgagc tggaagcggg 6480
aggcagaggg gctggccggc tgggtcagac acagtgaag cagaggctgt gacttggggc 6540
ctcagaacct tcacctga gctgccacc caggatctgg gttccctct tggggggccc 6600
cagggaacaa gtcacctgc cttgcatag gggagccct cagctatgt cagaaggttc 6660
tgctctgcc ctctccct ctaggtgctc agctctcca gccactagt cagatgtgag 6720
gctgcccag accctgggca gggtcatttc tgtccactga ctttgggat gggagatgag 6780
ctcttgccc ctgagagtcc aagggtggt gtggtgaaac ccgcacagg tggaagtggg 6840
catccctgc ccaggggagc cccagggac tctggtcact gggcttgccg ctggcatgct 6900
cagtctcca gcactactg acaccagcat ctactgacac caacattac aaacaccgac 6960
attgaccgac accgacatt accgacactg acattacca aactgttta ccaactga 7020
catctactga cactggcatc taccaact gacattacc gactgaca ttaccaaca 7080
ctattacca aactgacat ctactgacat tggcatctac caacaccaac attaccgac 7140
accaacatt accaactg aaattaccg acaccgacat ttaccgacac cgtttacca 7200
caccgacgt taccgacac gacattacc gactgata ttaccaaca ctgacatcta 7260
ctgacgtgg catctactga caccgatgcc agcatctacc aacaccgaca ttaccaaca 7320
ctgacattta ctgactga tatctactga cactggcatc tactgacac aacattacc 7380
aacaccagca tctaccaaca ccgacattta ccaaccag cattacca caccgatgtt 7440
taccaacgcc gacgtttacc gacccagca tctaccaaca ctgacattta ccgacaccga 7500

catttaccga cactgacatt tactgacact gacatctact gatactggca tctaccgaca 7560
ctgatattta ccaacgccag catctactga cactgatgtt taccaacacc gacatttacg 7620
agcaccgaca ttactgaca ccaatattta ctgacatcaa catttagcca tgtgatgggg 7680
gccggcttgg gggcaggcct tgctcttggc actggggatg ctgcagagac cagacagact 7740
catgggggtca tggacttctg ctctctctcc agcctcatgt actggacaga ctggggagag 7800
aacctaaaaa tcgagtgtgc caacttggat gggcaggagc ggcgtgtgct ggtcaatgcc 7860
tccctcgggt ggcccaacgg cctggccctg gacctgcagg aggggaagct ctactgggga 7920
gacgccaaga cagacaagat cgaggtaggg ctctgtgga catgtttgat ccaggaggcc 7980
aggcccagcc acccctgca gccagatgta cgtattggcg aggcaccgat gggcgcctgt 8040
gctctgctat ttggccacat ggaatgctg agaaaatagt tacaatactt tctgacaaa 8100
acgccttgag agggtagcgc tatacaacgt cctgtgggta cgtaagatgt tatcattcgg 8160
ccaggtgcct gtagacacag ctacttgag actgaggtgg gaggatcgct ggagtccaag 8220
agtttaggc cagcccgggc aaaggggaca caggaatcct ctgcactgct ttgccactt 8280
actgtgagat ttaaattatt tcacaataca aaattaagac aaaaagtaa tcacatatcc 8340
actgccctgc ttaagacaga aaacatgggt gttgtgaag ccagaggcag ctgctgcct 8400
gagtttggtg attggttct aagcagtga aggcagttt gttttccat agatgtctgt 8460
tctcccttg ctgggtgcag cctcgcctg ctgctgtgt cggtttcag tggcctcgtc 8520
ccgtggacgc agcctcggcc tgccgctgtg gtcgggttc agtggcctcg tccgtggac 8580
gcagcctcg cctgctgctg tggcgggtt tcagtggcct cgtcccgtgg acgcagcctc 8640
gccctgccgc tgtggtcggg tttagtggc ctgctcccgt ggacgcagcc tcgccctgcc 8700
gctgtggtcg ggttcagtg gcctcgtccc atgggcgtgc ttggcagct tttgctcac 8760
ctgtggagcc tctctgagc tttttgtt gtgtttgtt ttgtttgat ttgtttgat 8820
tgtttgtt ttgtgtcgtt gttgttggc aggcctggagt gcagtggcgc gatctcagct 8880

cactgaaacc tctgcctcct tgggttcattg ccattctcct gcctcagcct cccacatagc 8940
tgggattaca agtgcccgcc accacgcctg gctaaattt gtatttttag tagacagggg 9000
gtttcaccat gttggtcagg ctggtctgga actcctggtc tcacatgac cactgcctc 9060
ggcctcccaa agtgttggga ttacaggcgt gagccaccgc gccagcctc tgttgagcat 9120
atittgaggt tctcttggtg ccagtatat gtacatgtgt ccccatcgca ccatcgtcac 9180
ccattgaggt gacattgggtg cctctcctcg ggggtggatgc ctccctctgt ttccagcaac 9240
ttctgaagga ttttctgag ctgcatcagt ccttgttgac gtcaccatcg gggtcacctt 9300
tgctctctc agggctccca ggggaggccc gaatcaggca gcttcaggg cagggcagga 9360
tggagaacac gagtgtgtgt ctgtgttgca ggattcaga ccctgcttct gagcgggagg 9420
agtctcagca cttcagggt ggggaaccca gggatggggg aggtgagtg gacgcccttc 9480
ccacgaaaac ctaggagct gcaggtgtgg ccatttctg ctggagctcc ttgtaaatgt 9540
ttgtttttg gcaaggccca tgtttgcggg ccgtgagga tgattgcct tcacgcatcc 9600
ccgtacccg tgggagcagg tcagggactc gcgtgtctgt ggcacaccag gcctgtgaca 9660
ggcgttggtc catgtactgt ctacagcagt gttttctga gacagggtct cgctcgctca 9720
cccaggcgag agtgcagtgg cgcaatcacg gctcgctgta gcctcaatct ccctgggctc 9780
aggtgacct cctgcctcac cctctgagta gctgggacta cagacacata ccaccacacc 9840
cagctagttt ttgtgtattt tttgtggggg gagatggggg ttgctgttg tgcccaagct 9900
gatctcaaac tctgaggca caagcgtacc acctgcctcg gcctcccaa gtgctgggat 9960
gacaggcatc agccgtcaca cgcagctcaa tgattttatt gtggtaaaat aaacatagca 10020
caaaattgat gattttaacc attttaagt gaacagtca ggctgggcgt ggtggcttat 10080
gcttgaatc ccagtacttt gagaggctga ggtgggcaga tcacctgagg tcaggagttt 10140
gagaccagcc tggccaacat gatgaaatcc agtctctact aaaaatacaa aaattagccg 10200
ggcatggttg caggtgcctg taatcccagc tactcgggag gctgaggcag gagaatcgct 10260

tgagccccggg aggtggaggt tgcagtgatc tgagatcatg ccactgcact ccaatctgtg 10320
tgacagagca agactctgtc ttgaaaaata aataaataaa aaaaatttta aaaagtgaac 10380
aatcagggc attagtatg aggacaatgt ggtgcaggta tctctgtac tatctacttc 10440
tagaacactt tctctgtccc tgaaggaaac cccatgccca ccggcactca cgcccattct 10500
cccctctctc ccagcctctg tcaaccacta atctacttic tgtctctggg gggtcacttc 10560
ttctggacgt tttgtgtgac tggaaatcctg caatatgtgg tcctgcgtg tggcttcttt 10620
ccatagcatt gtgtttcca gattcaccca cacattgtcg cacgttatca gaatctcatt 10680
cctgactggg tgcagtgggt taggcctgta atcctaacat tctgggaggc caaggcggga 10740
cgatcacttg aggcaggagt ttgagaccag cctggccagc ctagcaagac cccagctacc 10800
aaaaaatttt aaaagttaac tgaacgtggt ggtggggggc acttgtggtt cccagctacc 10860
tgaggaggctg aggtgggagg atcgcttaag cccaggaggt caaggctgca gtgagctatg 10920
atcgaccac tgactccag cctggacaac agagcaagac cctgtctgaa aaaaaaaca 10980
aaaaaaaaag ttctttctt tttgtggctg gatgacatcc cattgtatgg ccacagcaca 11040
ttttgttgt ctgtttatcg ggtggggggc agtgggttcc acctttgtc tctgtgaat 11100
aatgtctgtg tgaacattg aattcaagt tttgttgaa cacctgtgt gaattattg 11160
gatatatgtg taggggtagg attgtgagt cctatggtaa tgtaggtt gacttactga 11220
ggaaccatta aactgtttc aacagtggct gcgccgttct gcatccccac cggcagtgtg 11280
tgagggttct gactttacct cctcaciaac gcttctttc catttaaaaa aatattcagc 11340
caggtgtctt ggctcacgcc tgtaatccca gcaatttggg aggccgtggc gggcggtatca 11400
cctgaggica ggagttcgag acgagcctgg ccaacatggt gtaaccccat ctctacaaa 11460
aatataaaaa ttagccgggt gtggcagcgg gcgcctgtaa tcccagctac ttgggaggct 11520
gaggcaggag aatcactiga acccgggagg cagaggttgc agtgagcaa gatcgcgcca 11580
ctacactcca gcctgggtga caagagtga actccatcta aaataaaaca aaaataaaaa 11640

tccagcacac actcaccctg tctgtgcacc tgttttgtg tccgtaagtg ggtatttact 32640
caccttacga gtgagccact gtgggaattc agggaggtgg cgagtgacc acccctggag 32700
ggataatgtg gtggcagggg tcgaggggtc cgccctccc tgcttctgc gcgtggcttt 32760
ctccaggacg gggagggctg agctgaagag gtggggacag ttgcgtccc ccgccacca 32820
ctgtctgcg gtgagagcag actcactgag cctgcccttc tccctgtgc ctccagcta 32880
catctactgg accgagtggg gcggcaagcc gaggatcgtg cgggccttca tggacgggac 32940
caactgcatg acgctgggtg acaaggtggg ccgggccaac gacctacca ttgactacgc 33000
tgaccagcgc ctctactgga ccgacctgga caccaacatg atcgagtcgt ccaacatgct 33060
gggtgagggc cgggctgggg cttctggtc atggagggcg gggcagccgg gcgttgcca 33120
ctcccagcc tcgccgacg tacctgtgg cctgcaagt ccccaacctg gcaggagctg 33180
tggccacacc cagactgcc cagcagctc accctctgt gtgggagttg tccccgtcca 33240
cccctgggtg ctttctgc agttatgtc ggagaggctc tggtagacgc tgttctgt 33300
gcacctgctg ggcactaggt ccagctaatt ccctgtgcca ggactctaatttaccctaa 33360
cacacatggt ggttttcatt gctggggaag ctgaggcctg agcacatgac ttgccttagg 33420
tcacatagct ggtgagttca ggtccccca gagataccag ggccagcact cgatccccac 33480
ccagccctga accccacat gtgtgggat tgtgtggga gtgtccacac gcctgggacc 33540
ccagggctgg tgctctcctc tcttttcc agatcatgag aatgaggctc agggaagttt 33600
gaaaaaaacc tatcccaagt cacacagcaa caggagcagg attgaaccc agaaaagggg 33660
accgcacact ctgttctgt agagtagtta gctgtcctgg gtgatatggc aggtgacagg 33720
ggcaactgtg ctaacaaag gaacccccat ccccctgcc aagtgggag actagaaggt 33780
caggggcaga agctctgaag ggccaggtgc agtggtgac acctctaate ccagcacttt 33840
gtgaggccaa ggcgggcaga tgattgagc ccaggagttc aagatcagcc tgggtaatgt 33900
agtgagacgc catctctaca aaaaaattt taaaaatta gctgggcatg gtggttcag 33960

cctgtagtcc aagctacttg ggaggctcag gtgggaggat tgcttgagcc caggaggttg 34020
agggtgtggt gagctgtgat catgccactg cactccagcc tgggcaatag agtgagaccg 34080
tctccaaaa aaaaaaaga agaagaaaa gaagctctga ggctccaagt ccccaggcac 34140
cccttggtt gagggcagac aaggaggag agggtcacct gggcagccct gactttgtc 34200
ccctggcaaa gggaccttca gtgaccttg ccctaggaga gcctctgagc acgtcagcca 34260
tgtgaaccg ctgaggaagg gcagcaagaa ttggcttct gacctctgcc tctctactc 34320
gccatctgca ctgggtgtgg ttgtgcccac ttacagatg aggaggctgg ggcatcgacc 34380
agctgaatgc ctgtcccag gtactgcgta ggcagagctg gcagtgaac cccgtgtcct 34440
ggttgtcgct ggggggtgggc tgcaccctga ctgtgaggc cagtagcaag gttgcacgt 34500
gacttctga ccgtcaccca gctctgcagc acatccctg acccagctca tccaggccgc 34560
atgcaaacct gttgccaggc gagaaaccag tcaccgcaca gctgtggtg cctgaaatga 34620
ttaagctcat taatacccc ggagtgagga cagactcaga tgaaaaccag caaaagccct 34680
ggaaactcat gtgacctgc caatgagggc ggccatgtgc attgcagcct ggccgtcact 34740
cctcggtacg tgttttgac ttaaaccgtc cggatgtta ctgagtgtt gattaataac 34800
atggaaggcc tggctcatt gctgtgggag tgaaggatgc acagccaggc ctgacatgat 34860
gagaacaaga acctggagtc tcgctgcctg ggtggtaatc ctggccctgc cacttagcaa 34920
ctgtgtgact gtagccaggt cacttaattt tgctagatcc tgcctgcgtc tcagtggatc 34980
ttgttggtt tccaagggtg ccaaacactt taaggcattc atgtggtcgc taggctgcag 35040
ggttgaacct tggctaccc cgcaggggcg cgtgtgctct gtggccctggc tgtgccttg 35100
ctgacaccgt gcccgtgtgt gttcatgcag gtcaggagcg ggtcgtgatt gccgacgatc 35160
tccgcaccc gttcggctcg acgcagtaca gcgattatat ctactggaca gactggaatc 35220
tgcacagcat tgagcgggcc gacaagacta gcggccggaa ccgcaccctc atccagggcc 35280
acctggactt cgtgatggac atcctggtgt tccactctc ccgccaggat ggcctcaatg 35340

actgtatgca caacaacggg cagtgtgggc agctgtgcct tgccatcccc ggcggccacc 35400
gctgcggctg cgctcacac tacaccctgg accccagcag ccgcaactgc agccgtaagt 35460
gcctcatggt cccccgcacc tcactccctc gttagatcag gctggttctg ggagctgacg 35520
ctgaaaggag cttctcatct ggggttcctg ggtgtacata gatggttggg taggttgtgc 35580
actgcacaag ctgcatgatg ctacctgggg gtccaggctc aggctggatg gacttgttgc 35640
ttcatcagga catagataaa tggccaaaac tcctcagctg gaaggctctg ggcaggatct 35700
ttgggtgtga aaaccagtca cagggaagg gtgctgctc atactgccag cacagtgtgc 35760
agtgccttcc atagcgctcg ttactcctc aagcctggag ggtggggagt agcatgttcc 35820
catttcacgt acaaggaacc cgatgcacag agaggtgtgg caacccatcc aaggccatac 35880
aactgggggtg ggttgagccg ggggtgactg tggcaggctg gctcaagagt ccttgctcct 35940
gaacccttgc caggcagcct ggcatcagct cggggaattt ttgccctgac ccttgaagc 36000
aagtgggcct ctttgtctc atgtcagtga tgagaagagt gactttccta tggcccctct 36060
ggagtacagg tgttctctgt tggcgggctc ttccccatg acatcagcag cgagctggtt 36120
atgattccct acgcagaact tgatagtta taaagctctt tgtcatccag gcccgttgg 36180
agtctcacgc agacctggtc gcaggcgggg ctggtcttgc ctgtcccagc tgcattgatg 36240
gggaacttga ggcttgcaaa ggtaagggg ctgttcgagg ccacagctgg caggagatgg 36300
gcctgggcca gagtctggga ctcccatgc ctgggctgtc ttggtctctg ttgtcacca 36360
tcctccctg gggccatgac ctagagagc caaatggagg tgcaggtaac ccacggcaag 36420
gagggttgc catgactcag agtccccgtc ctgtggccgg cagtacctgg tgcaacgact 36480
tggtttcag accagccact gtagcccgtc gacggtgcgc tcgaagtgcc acagcttctg 36540
aagccaggca ggactcaggc caggagactc tgttagctgt tgagaggag aggccaacgg 36600
atgttctggt tctgctagag agctggttct tcggatcctg gtaccagtgc actgagagga 36660
ggcccagctt gattctgggg ctgccttgtg gtggcatgtg ctgctcactg acaccctcga 36720

gccagggcac agccctcaaa atcccatca atgacatga agaaaagaga ggaacctggg 38160
aaatagcaaa gtgccttttg cacattaaat ggtagctat atcccacaat actgtgcatt 38220
cgtaaacgtt aatgtgcaa taaatacggc actcacctt gggaagatct ggagttggct 38280
tatgagtgtg gaaggggtga gcgcatgagt tttgtgaaa cactggaagg aggattgtgg 38340
gaaatcaaat ggaaagtct caccacaggc gtggagaaga gtgggtcatg gcccagcag 38400
tgagcccagg gaggtcagag acggaggtgt gtgtgtgggt gtgacctgc gcagttccct 38460
gccggctgta gtttttgca ttcgcttaat gttctcgtg gaggaaattg tgcattgagca 38520
aatgtgaaac cgtgctgtgc tcaaatgtc ctaatacatc attgcattgg aacagattgg 38580
ctttntttt tttttttt tttttttt ttgaaatgg agtctcactc tgcaccagc 38640
ctggagtga gtggcatgat ctggctcac tgcaacctt gcctctatg ttcaagtga 38700
tttctgcct cagcctcctg agtaactggg attacagggc atgagccacc gcggccggcc 38760
agatttgcatt tttgaaaca actgctaggg tgggcgggt ggctcacacc tgtaatcca 38820
gcactgtggg aggccgaggc aggtggatca cctgaggtca ggggttcgag accagcctgg 38880
ccaacatggt gaaaccccgct ctctactgaa tatacaaaa tcagctgggt gtggtggcgg 38940
gtgcctgtaa tcccagctac tcaggaggct gaggcaggag aattgctga acccaggagg 39000
cagaggttgc ggtgagccga gatcacacca ttgactcca gcctgggcaa caagagcaaa 39060
actccatctc aaaaaataaa aaatagaaaa acaagtgtg tagcggaaat gagcacttg 39120
cggagtcagg ctgtgtggc ctgtccaca aatgatgtg tcacgggtgc ctgagccca 39180
cctggagtct gcagcatggg gcacaacagg ttcatatg tagaattca ggacaggcct 39240
ggctcctaag cagccttctt ttacaaaaac tgcagagccc gcctgtatc tagcacttg 39300
ggaggccgaa gtgggtgat cacgaggtca ggagtcaag accagcctgg ccaacatggt 39360
gaaaccccat ctctactaaa tatagaaaa ttgctgggt gtggtggcac gcgcctgtg 39420
tcccagctac tcgggaggct gaggcagaat tgctgaacc tgggaggtgg aggttcagg 39480

gatctgagac catgtcattg cactccagcc tgggcaacag agcgagacgc catctcaaaa 39540
aaaaaaaaacc tacagagcca cacggcctct ttctccaccg agtggttggtg tgggagcttg 39600
tgttattgtg gtgaaatctt ggtactttct tgaggcagag agaggctgag cgcctggaga 39660
gactttcaca tgggtcgcca tgtccgccgt cggtttcgct gttgtgctcc ccatctgaag 39720
gctggtgccg tccagacagg ctggacgccc cttccacca gatccttct cccgcagcag 39780
tttctagta cgttgtactg tgaggtctgt gtccttggt gatggcaaaa gtcagccgaa 39840
ttgaaattca gagccatgcc tggctccctg gagcttctct cctgggcagc tgtgatcatt 39900
gcctctgctg tgggtgagggt ggtggaaatg gattccttc atcttgcttg ctacaggtga 39960
ctgtcacgtg gaggccttg gagagaggga cgtgttaatt gatggatgtg gtccccatgc 40020
tgagaaagct cctgggcgta cattgccta gagtttcatt ggagctgcgt tctttatgg 40080
tgtctgctag gcagaagtga tgaagacttg gaagaaaacc cagaaggttt tccacttaat 40140
ttggaaaatg tgcctttccc ctctgtgtc ttgtctaag gtccagcctc ctgcagcctc 40200
cccgtctgtg ggactctggc ttgattctt tattaggagt cccctgctc ccccaaaaga 40260
tggtgtctaa attatcatcc aattggccga ggttttgtt tctattaatt gttttattt 40320
tttattgttg taaatttata taacataaaa ttgccattt taattgttt gttattgtg 40380
ttttgagac aggtctcac cccagtccc aggtggagt gcagtgggtc gatcatggct 40440
cactgcagcc tcagcctcca gggctccagt gatccttca cctcagcctc tctagtagcc 40500
gggactacag gcatacacta ccacatctgg ctgattttt gtatttttt ttattgtag 40560
agaccgcta tgtgcccag gctggtctca actcctggac tcaagccatc ctcccacctc 40620
accctccaa agtgctggga ttacaggcat gagccacaac acccagccat ttaattttt 40680
tttttttt ttgagatgga gtctactct atgcccagg ctggagtga gtggcgtggt 40740
atcaactcac tgcaacctct gcctccagg ttcaagegac tctcctgcct cagcctctc 40800
ccgagtagct gggattacag gtgcccata ctatgcctgg ctaattttt tatttttag 40860

ggagtgtctt ctctcgggct tgttgactgt gcccggttt ccgcagtca ctggtgcaca 36780
cataggcaca tagcaaaccg cacacacagt cgtgggtatg agtttacta cattccacca 36840
ccagtgttca ctaccattac ctgccttccg tcttaagtgt tcatcatta aaaataaatt 36900
tattgggctg gacgcgggtg ctcagtactg ttatcccagc actttgggag gctgaggcgg 36960
gcagatcacc tgaggtcagg agttcaagac cagcctggcc aatatggtga aactccatct 37020
ctactaaaaa taaaaaatta gctgggcatg gtggggcatg cctataatcc cagctactca 37080
ggaggctgag gcaggagaat ggcgtgaacc cgagaggcag agcttacagt gagcccagat 37140
agcaccactg cagtccagcg tgggcaacag tgcgagactc catctcaaaa aaaaaataaa 37200
taaataaaag aaaaaataat ttatgatcta ttcaaaaat aacacatgta ctttgaaaca 37260
gcagagacac atatgacacg gagaatgaaa ttccccatag cgcaccccca agagacagcc 37320
ctgtgcccc cgtctttccc gtggacctcc agcggggcag atgtgagcc gcctgttgc 37380
gagtggcatg ctatcccgct ctccagctcc tctgtggctt acagacacc acctgcagcc 37440
ctgtcttgc ctctctagc gccaccacc ttctgtctgt tcagccagaa atctgccatc 37500
agtgggatga tcccgacga ccagcacagc ccggtatca tctgcccct gcatggactg 37560
aggaacgtca aagccatga ctatgacca ctggacaagt tcactactg ggtggatggg 37620
cgccagaaca tcaagcgagc caaggacgac gggaccagg caggtgccct gtgggaaggg 37680
tgcggggtgt gcttccaag gcgtcctct tctgtgttc caggctgtg cccctgtct 37740
tagcagaggg aggaaacaga ggtggctct gggatgaatga tgactgggc ttcgattatg 37800
tagtcacagg gtatgacct gagatgcgtg gaaccccgag actgtgatta tatgtagaaa 37860
ctgggttcc ccgtgttta agtagtcag gtggggtcag accccacagg actttgtct 37920
ttcaagaaa gaaaatggtc gtgtgtcatg caggggtagt tggtaggt taatccaggt 37980
ttatccttta tttgtggga actgtacagt cattctgct acaatgctgt atatgctct 38040
ctgaaagaca cctatgcaa atcgcacagt aaaaatgaca caactcatag ggaaagcggg 38100

cagagacggg gttcaccat gtggccagg ctggtcttga actcctaacc tggatgccg 40920
cccgctcgg cctcccaaaa tgctgagatt acagggtga gccaccgtgc ccggccttt 40980
tttgttttg agacagggtc ttgccctgc acccagactg gagtgaatg gtgggctctt 41040
ggctcactgc agcctccgc tcccaggctc aagtgtga cctccacacc tggtaactg 41100
tattttatgt agagacagat ttcaccatgt tgcccaggct gggcttgaat tggactcaag 41160
cagtccaccc acctcagcct cccaaagtgc tgagattaca ggcgcgagcc accgcacca 41220
gcccatttta cctattctgc agttgacagt tcagtggcat tcagtcagtt cacgaggtaa 41280
ccatcactgc cattcatctc cagactactt caccttctcg gcagatgtcc gaaactgtcc 41340
gcattgaaca cactcctcat ctccctctga cagccacat tctactttgt atctctctct 41400
gccttctcta ggtacctcat gtaagtggaa ttataccaat attgccctt gtgtgactgg 41460
cttctttcat gtgacatggt gtctcaagg ttcatctgtg ttatagcctg tgcagaatt 41520
tccttcctta aagcctgaat aataaccgt tgtaaaggct ggcgcggtg gctcacacc 41580
tctaatacca gcatttggg agtccgaggt ggcagatca cttgaggtca ggagttgag 41640
accagcctgg ccaacatagt gaaaccctgg ctctactaaa agtataaaaat tagctgggtg 41700
tggtggcgcg cacctgtaat cccagttact caggaggctg aggcaggaga atcgcttga 41760
ccggggaggc agaggttgca atgaaccaag attgtgcctc tgcagtccag cctgggtaac 41820
agagtgaac ttctgtctc aaaaaaaaaa aaaatcatcg gatgatgga cggaccactt 41880
ctgttattt atccatccac ggtgctagg ttctccac ctttggtgt cgtgaataag 41940
gccactatga acatttctt ccgtggtgaa ggtttgtac tagtgaggaa aaggcgtgtt 42000
tgtggtgtg cataggattc tggaagaaa gttgacta accataagta ttgtactac 42060
attaaaatga aagtcaggg gccggcgcg gtggctcac cctgtaatcc cagcactttg 42120
ggaggccagg gcggcgcat catgaggtca ggagatcaag accatcctgg ccaacatggt 42180
gaaacccgt ctctactaaa aatacaaaa aactagccag gtgtggtggc gggcacctgt 42240

ttcaagtgaat cctcttgccct cagcctcctg agtagctggg attacaggtg cagccacca 51960
taccagcta atttttgtat ttttagtaga gacgggggtt caccatgtg gtcaggctgg 52020
tctcgaactc ctgacctcgt gatccgcca cctcggcctc ccaaagtgt gggattatag 52080
gcgtgagcca ctgtgcctgg ccattaggtg tgtttatca cccagcatca tgcagttat 52140
cttggtgaat gttctgtgia ctctgaaaa gaatgtggat tctgctgtg ttgggtggag 52200
tgttccagaa acatcaatta gatccagtg gtaaatagt ctcacaggt tgtctctatc 52260
cttcttctt gactgcctgc ttgagctgc agttattgac aggggtgtgg agtctcaac 52320
tctaagtgtg gatgtgtta tttctcctag tagttctatc ttttctctc ctctaccct 52380
tgatcctctt cccccctag ggcttctgg tgttgggtg gggagagtgg ggtagtgaag 52440
aacctggact ttaggccaa agaggccagg gttcaaatcc tggctctgc acttccagt 52500
tgagtgacc tggctgggtc ctgaatctt gtgagcctc acttctctt ctgtgaaatt 52560
gagagcactt acctggcagg ctgtcatggg catcaagtaa cagggcactc cacctggacc 52620
ctgacacgtg atgcacagga atgccagctg ctatgccatg ggtgtggcag tagtaataaa 52680
gtgaccatct gtatcctcac cacagtgaag cctgtccagg gctttctctc ctatgcccc 52740
atgcctccag gtggccttgg atcctgttgg ttctgtgctc tgctcagcga cttttctcc 52800
gtgggagttc ctgggggttc agcttcatcc tacagacagc agcacacact ggctgtgcac 52860
cctttttt tttttttt ttttttga gatggagtct cgctttttc gcgcaggctg 52920
aagtgcagtg gtgtgatctt ggctcactgc aacctctacc tctgggttc aagtattt 52980
cctgcctcac cctcccaagt agctgggatt acaggctccc accaccacgc ccggctaatt 53040
ttgtatttt cagtagagat ggtgtttcac catgttggc aggatgtct tgaactcctg 53100
acctcaggtg atccgcccac ctcagcctcc caaagtgcag ggattacagg cgtgagccac 53160
cacacccgga gtgccggtg ttttagcag ttgtctgt tcttgagag actggctcct 53220
gccagggagc tcggggagta gggccgcggg gtgctgcctc acacctgag ttggccgta 53280

atttgtgtg ccagcgcgtg gtgtgccagc gctatgcggg ggccaacggg cccttcccgc 60240
acgagtatgt cagcgggacc ccgcacgtgc cctcaattt catagccccg ggcgggtccc 60300
agcatggccc cttcacaggt aaggagcctg agatatggaa tgatctggag gaggcaggag 60360
agtagtctgg gcagctttgg ggagtggagc agggatgtgc taccacaggc cctcttcac 60420
atgtggcaga cattgctaatt cgatcacagc attcagcctt tcccactgag cctgtgcttg 60480
gcatcagaat cctcaacac agaggcctgc atggctgtag caaccaccc ttggcactg 60540
taggtgtgga gaaagctcct tggacttgac cttcatattc tagtaggaca tgtgctgtgt 60600
tgtccacaaa tctcatgta ccctagaaat gaatgtgggg gcggctgggc tctctccaga 60660
gctgaaggaa tcaactgtga ccatacagca gctttgtctt gactgcagct gggatttgtg 60720
gctgagcagt tacaattcct acgtggccca ggcaccagga acgcaggctg tgtttgtaga 60780
tggctgggca gccgcaccgc agagctgcac catgctggtt tgtatcacat gggtgacat 60840
ggtatgtcta agaaggtgga gtccctgtga ggtctgcagg tgccccaca gctccaggcc 60900
accttgagga ttgcctctgc ctgcccagcc ctgagttccc tctcccctgt cctgtccac 60960
tgtcacccca agccggcctc attgggagcc tgttgatgg cagggtatag atgtaacctg 61020
attctctctg gggagcgggg ttatctggct tctcaagagc tctaggagc ccacagtgg 61080
ggcaccatca cagtgcagc agccccaga gaacgcggcc ctgtctgttc ctggcgtgct 61140
ctgtgctgcc ccgcctgggt tccctgcccc agtcgcaggc cccttgagg aggtaccatg 61200
tgtctcccgt ttacagatg agccccgggg agtcactct agtagtgcc agagaggcct 61260
gcggctcagg gagcggggca cattccaac aggacacacc gccctggtct gactctctg 61320
ggtagtggga gcagaggaga gcgccctatg tctgtggggc ggcttggtg agcctggaag 61380
ccacctgacc tccccgtcc cttccctgcc aggcacgca tgcggaaagt ccatgatgag 61440
ctccgtgagc ctgatggggg gccggggcgg ggtgccctc tacgaccgga accacgtcac 61500
aggggcctcg tccagcagct cgtccagcac gaaggccacg ctgtaccgc cggtagggg 61560

tggactcgag cgatcctcct gcctcggcct ccccagggtgc tgggattaca ggcgtgagcc 63000
accgtgcctg gcctggggta tigtcttctt atggcacctg actgtggtgg gccctgggaa 63060
ggaagtagca gaagaggggt ctctcttggtt tcctggacag taactgagtg ttctggaggc 63120
cccagggcct ggctttgttt agggacaaag ggaactggta accagaagcc gagagttaa 63180
acacccactg cccttcttcc ctgctcctgc tgctgcaacc cagcttaacc agccaggagt 63240
gctaggaacc caagcagggc ccccgagcac acagcaggca gctcacgaat tctctttcc 63300
tgttctccct tgggagctgg gaggatctta atcaggcaat aagagatggc actgagcagc 63360
cagctaattt tttaaatcac ttattgttt aaccatatga ctcacccact taaaaaagg 63420
tacagttcag tgggttttag tgtattcaca gatgtgtgca accctcacca cagttaattt 63480
tagaacattt tcctgccctt aaaagaaact ctgcatgaag ccagctgttt ttaaattagc 63540
aaagttattt tgcatccttt aaatataigt tcatgttaca aaattcaaaa gatacagaag 63600
agtctgcagt ccaaagagac tccgccccca tgacgccaag caggactccc tgggaggcat 63660
ggcctcctgc agtgtgtttc ttctatgtcc ccccaggggt catctgtaca tatgcaagca 63720
tacaagagcg tggactttgt ttccaagcc agaagataat ttagattta tgtgcagttg 63780
tgagaaagag cacagacca ttatcctct gcctggttc cccagtgct gcctgccatc 63840
ttgcatgact tccattccta tcataagcaa gacactgata acgattctt caccttattc 63900
agattgacat aagtgtttt tttgttct tgagacaaac ttctctgtc acccagtggg 63960
agtgcagtg cacaatcaca gctcactgca gcctcaaact cctgggctca agcgattctc 64020
ctgcctcagt cccctcaagt agctcagatg gcaggtgtgc accatcatgc caggctaatt 64080
tttaaattt ttgtggagggt gaggcctcac taaatttct gggctagtct tgaactcctg 64140
agctaaagt atcctcctgc ctacgcctcc caaagtggta ggattacagg catgagccac 64200
tgcgcctggg ctgacatatg tgtttctga agcccgaag atagcatctg aagagtcaac 64260
attgagcctt gccttttgc tctaattgat tataaaagct gctgttctga gcatttcgga 64320

gggggcttgg gctccaaact ctgactgtgt gtttgagtcc ggctgtggaa acctagccat 68520
tgagatgccc cctcttggtg gctctgtcct cttaggatgg gacaagtctg tgaaggctgc 68580
tgcagcacc accgtagacc cctaactgtg tgacgtcacc aggatgggtcc gggctgtctca 68640
cttgccacag tggcctgttt gagcccggga agccaacggg gctgtcagc tggacaccag 68700
ccccccgagc tgcccatgtt ggggtcacag gcccacctc cctggttggg gaggggcaac 68760
tgagagtgtg gagaggtggg acccaggtgt gctggtctcc gcaggggctg gatcagagcc 68820
tgggatgggc aggggtgagcc tctgacctt taaccagtg gtgtcaggca acgtggccca 68880
cccgccagcc gcaccaggcc ccacccccgc aggtgaaggg gtgggatagg ctgggcctgg 68940
gccaggacac ctctggacca cgcatcctc attgcttggg tccctggagc agcagggcct 69000
cccgagtgtg gtgccgcctg ccacctagt gccattcca cgaactcca ggcttggtg 69060
gggagccgga actgcagcct ccatttcac cccactccgg gtcgggccac ctccctgatg 69120
cctcagtatt atatcaaact gtcacagtct gtccacagc ctacagacc actgtctcca 69180
gaatggtcac atccacactg ggcagcccag tctcgtagt tctctgtccc acctctgcc 69240
tttgcctatg cccgtctgc tctgggcccc ccgcggacac atctcccc cgcccgcgt 69300
ctgacctcac agcagctggg cccaagagg agtatcctgt cctgtgcac ttttcaac 69360
accgggtgtt ggctgcacct tcccacccat tgcaggcccc tctgtgacag gacgggggct 69420
cctaaacaca ccacagtcc gagtctgaac tcacacagt ggatgcggcg tttctgggcc 69480
acagtgggt gcaggtagcc tctgggagga tgggaggtca ggagccatct tgcgagtcag 69540
gttgcttgaa ctaggatgg aagtgtccg ggcccattg ttgctgtatt agcctgttct 69600
cacgtgcta ataaagacat acccaagact gggttaattg aaaggaaaga ggttaacgg 69660
actcacagtt ccacctgcct ggggtggcct cacaatcatg gtagaagaca aggaggagca 69720
agtcacatct tacatggctt cagggaacag acagcatgag aaccaagcga aaggggtttc 69780
cccttgtaaa accatcaagt ctagtgagat ttattcacta ccacgagaac agtatggggg 69840

ggctgctcag ctggacacca gcccccgag ctgcccattgt tggggtcaca ggccccacct 2940
ccctggttgg ggagggggcaa ctgagagtgt ggagaggtgg gaccagggtg tgctggtctc 3000
cgagggggct ggatcagagc ctgggatggg cagggtgagc ctctgacct ttaacccagt 3060
ggtgtcaggc aacgtggccc acccgccagc cgcaccaggc cccacccccg caggtgaagg 3120
ggtgggatag gctgggcctg ggccaggaca cctctggacc acgcattcct cattgcttgg 3180
gtccctggag cagcagggcc tcccagtggt ggtgccgcct gccacctagt ggccatttcc 3240
acgaactccc aggcctggct ggggagccgg aactgcagcc tccattcca cccactccg 3300
ggtcgggcca cctccctgat gcctcagtat tatatcaaac tgcacagtc tgtccacag 3360
ccttacagac cactgtctcc agaattgtca catccacact gggcagccca gtctcgctag 3420
ttctcgtcc cacctcctgc ctttctcat gcccgctctg ctctggggcc accgaggaca 3480
catcttcccc ccgcccgcg tctgacctca cagcagctgg gcccgaagag gattatcctg 3540
tctgtctga cttttctcaa caccgggtgt tggctgcacc ttccacca ttgcaggccc 3600
ctctgtgaca ggacgggggc tctaaacac accacagttc cgagtctgaa ctacacagt 3660
gggatgcggc gtttctgggc cacagtggg tgcaggtagc ctctgggagg atgggaggtc 3720
aggagccatc ttgcgagtca ggttgcttga actcaggatg gaagtgttc gggccattg 3780
gttctgtat tagcctgttc tcacgtctgt aataaagaca taccaagac tgggtaattg 3840
taaaggaaag aggtttaacg gactcacagt tccacctgcc tgggtggcc tcacaatcat 3900
ggtagaagac aaggaggagc aagtcacatc ttacatggct tcagggaaca gacagcatga 3960
gaaccaagcg aaaggggttt cccctgttaa aaccatcaag tctagtga tttattcact 4020
accacgagaa cagtatgggg ggaaccaccc ccatgattca atcatctccc actgggtccc 4080
tcccacagca cgtgggaatt atgggagtag aattcaagat gagattggg tggggacaca 4140
gccaaacct atcggttgcc aacatttaca gtaacagtgt taggtgaaca gtgtccagt 4200
ctctgtttt gtcggacact gttttagca ccttcaggc agaattcat gtatcctca 4260

cttfcgaaat ggggtactatt tcatccccac tttatcaat gagaaactaa agctcgaaga 4320
gggtcaagtaa gttcctggcc aaggtcagct agcaggctct agaggcctcg ttctccttag 4380
aggcagcctt gccaggggccc aggcttgga ggctgcaggg cagggtgcggg catgcccattg 4440
gtagagggtgg gaccattgag gctcagagag ggtaagtgat gagccctggc gacacagcgg 4500
gggtgggtcca gagtccggcc tgcattctt ggagctggcc agtggacagg cctttcccgt 4560
tcacagcccc ggggctgctg tgcccaccag ggcggatgtg cctaccgaat cccactcctc 4620
tgtgtgtgtc ccttcaggc cctacatcat tcgaggaatg gcgccccga cgacgccctg 4680
cagcaccgac gtgtgtgaca gcgactacag cgccagccgc tggaaggcca gcaagtacta 4740
cctggatttg aactcggact cagacccta tccaccccc cccacgccc acagccagta 4800
cctgtcggcg gaggacagct gcccgcctc gcccgcacc gagaggaget acttccatct 4860
cttccgccc cctccgtccc cctgcacgga ctcatcctga cctcggccgg gccactctgg 4920
cttctctgtg cccctgtaaa tagttttaa tatgaacaaa gaaaaaata tttttatga 4980
tttaaaaaat aaatataatt gggattttaa aaacatgaga aatgtgaact gtgatggggt 5040
gggcagggtt gggagaactt tgtacagtgg agaaatatt ataaacttaa tttgtaaaa 5100
cagaactgcc attcttctgt gccctgtgtg catttagatt gtgtgtccc gtggagggaa 5160
tgccgacccc cggaccacca tgagagtctt cctgcaccg ggcgtccctc tgtccggctc 5220
ctgcagggaa gggctggggc ctggggcaga ggtggatata tcccctggga tgcattccctg 5280
agctgcaggc cgggccggct ttatgtgcgt gtggcctgtg ccgtcagaaa gggccctggg 5340
cttcatcacg ctgttctgt tcttcttct cagattctta gtctttttt tttttttt 5400
tttttgaga cggagtctt ctctgtcatc caggctggag tgcagtggta caatctcagc 5460
tactgcaag ctccgactcc caggttcaag tgagtctct gcctcagcct cccagtagc 5520
tgggactaca ggtgcgcgcc accacacccg cccagctaatt tttgtattt ttagtagaga 5580
tggggtttca ccatgttggc caggatgatc tcgatctctt gacctcgtga tccgccacc 5640

aaaaaacaag acccaaccat ctcttgcata caagaaacac actttaccta taaaaacaca 7080
ctaggccagg tgtggtggct cacacctgta atcccagccc ttggggaggc ctgactggca 7140
gatcacctga ggccaggagt ttacagaccag ctgaccgac atggcaaaac cccatctctc 7200
ctaaaaatac aaaaaaaca aaaaaagaaa aaggctggaa gtagtgatgt gtgcctgtag 7260
ccccagctac ttgggagggt gaggcaggag aattgcttga atccgggaag tggaggttgc 7320
agttagccag gatggtgcc ctgcactcca gcctgggtga cagagcgaga ccctgtcata 7380
aaaaaaaaa gaaaagaaaa gaaaaacgag aaaaacaaac aaaaattag tagaagaaaa 7440
gaaataataa agatcagaac aggccaggct catgggcaca gtggctcaac tcctacctgc 7500
tcaggagttt gagaccagtc tggccaacat ggcaaaaccc catctctctt aaaaatatga 7560
aaaaaaaaa ataggctgga tgtggtgatg tgtgtgtgcc ttagcccca gctacttggg 7620
aggctgaggt gggagaatca cttagccca ggaagtggag gctgcagcga gtcatgaatg 7680
cacctgcac tctagctggg taactggagt gagattctgt ctcaaaaaag caaagaccag 7740
agcagaaata aatgaaatgg aaatgaagga aacaatgcaa aatgatacaa aaagttttt 7800
cgaaaagata acaaaatca acaaacctt agccagatta agaaaaaag agagaagacc 7860
caaataataa aaatccgaga ttaaaaagga gacattacca ctgataccac agaaattcaa 7920
aggatcatta gaggcaacta tgtgcaacta tatgctaag aactggaaa cctagaagaa 7980
ctgggtaaat ttctagacac atacaaccta tcaagattga accatgaaga aatccaaaac 8040
ctgaacaggc cgggcacggt ggcttacgcc tgtaatcca gcactttgga aggcctgaga 8100
tcaggagttc gagaccagcc tggccaacat ggtgaaaccc catctctact gaaaaaatat 8160
aaaaattagc cgggcgtggt ggcgggtgcc tctaagtca gccactcggg aggctgaggc 8220
aggaaaatca ctgaacctg ggaggcatag gttgcagcga gccgaggtt caccactgca 8280
ctccagcctt ggcgacagag ccagactcca tctcaaaaa attaaaataa caaaaacctg 8340
aacagacca taacaagtaa tgcgatgaaa actgtaataa aatgtttcc acaaagaaa 8400

gcccaggaac aaatggcttc actgctgaat ttaccaaac atttttttt tttgagacg 8460
gagtctcgct ctgtcgccca ggctggagtg cagtgggtga acctcggttc gctggtaact 8520
tatgectctc aggctgcaag tgattttct gcttcaggcc ccccgagtgg ctggaaatta 8580
gatggactt gtcaaacaag gcctggctaa attctatat ttccttcaag tagaagatgt 8640
gcttccaaca aaggttgggt tacggctggc ttctgaaaat cttggatttc aaggctcccc 8700
aaaag 8705

<210> 11

<211> 66933

<212> DNA

<213> Homo sapiens

<400> 11

tataatcaag cgcgttccgt ccagtcagggt gggaagatt tcgatatgct tcgtgatctg 60
ctcaagaacg ttgatcttaa agggctcgag cctgatgtac gtatttggct taccaaatac 120
agcaatagta atggctctca gtccccgtgg atggaggagc aaattcggga tgcctgggga 180
agcatgggtc taaaaaatgt tgtacgtgaa acggatgaag ttgtaaagg tcagatccgg 240
atgagaactg ttttgaaca ggccattgat caacgctctt caactgggtc ctggagaaat 300
gctctttcta ttgggaacc tgtctgcaat gaaatttcg atcgtctgat taaaccacgc 360
tgaggagatta gataatgaag cgtgcgcctg ttattccaaa acatacgctc aatactcaac 420
cggttgaaga tacttcgtta tcgacaccag ctgccccgat ggtggattcg ttaattgcgc 480
gcgtaggagt aatggctcgc ggtaatgcc ttactttgcc tgtatgtgt cgggatgtga 540
agtttactct tgaagtgtc cggggtgata gtgtgagaa gacctctcg gtatggtcag 600

taaataaaaa ttattaaaa cattcatcac agccagccta gtgggtgtcc catgtggctt 11700
tgccctgcat ttcctgata actaggatgc tgagcgtctt gtcccaggct tgccacacct 11760
cagcactttg agatacgtcg cacagtcccc atttgcgaac gagaaatgag gtttagggaa 11820
cagcagctgt gtcattgtcac acagcgagca gggggtctct gagccgtctg accccacagc 11880
cgaccaagct ccaatcctta ccgcctccta gtgtgtgga ttagccagc ggtgtccca 11940
cattttcag atgagaacac cgaagctcaa aacaggagcg tttgtccac attggataca 12000
cgatgtctgt ggttgggtcc tgaagtcact ttatatctca gtgtccaga ctggagtagg 12060
acaggggggt ctggggaatg gggaagggtgt ctgagtgaa aggaaggaat tccagattct 12120
ccatactgtc ctgggaagt tagaagactc agagggtctg gcaaagtcag acaaagcaag 12180
agaaatgcag tcaggaggaa gcggagctgt ccaggaacag gggggtcgca ggagctcacc 12240
cccaggaact acacttgctg gggccttctg gtcacaatga cgtgagcact gcgtgttgat 12300
taccacttt tttttttt ttgaggtgga gtctcgtctt ctgcccagt ctggagtga 12360
gtggcacgat ctgggtcac tgcaagctct gcctcccggg ttcattgcat tctcctgctt 12420
cagcctcccg cgtagctggg actacaggcg cctgccaccg cgcccggcta attttgtat 12480
tttagtaga gatgggattt cactacatta gccaggatgg tctgatctc ctgacctcat 12540
gatccgcccg tctggcctc ccaaagtgt gggattacag gcgtgagcca ccgcgcccg 12600
cccgatttc cacttaaga atctgtctgt acatcctcaa agccctatac acagtgtgg 12660
gttgctatag ggaatatgag gcttacaggc catggtgtg gacacacaga agggacggag 12720
gtcaggaggt agaagggcg agagaggaa caggcggagg tcacatcctt ggctttcaa 12780
atgggccagg gagagacacc ctctgagcat ggtaggacag gaaagcaaga ttggaacaca 12840
ttgagagcaa ccgaggtggc tgggcgtggt ggcttacgcc tgtaatcca acactttgga 12900
aagctgaggt ggggtgattg cttgaggcca ggagtcaag accagcctgg ccaacatggt 12960
gagaccccg ctctactaaa tataaaaaa ttagccaggc gtgatggtgc atacctgtaa 13020

tcccagctgc ttgggaggct gaggcaggag aattgcttaa acctgggagg cggaggttgc 13080
agtgagccga gatcccgcca ctgcactcca gcctgggcca cagagtgaga ctccatctca 13140
aaaaaaaaa aaaaaaaga taaaaagacc aaccgaggaa ttgaagtggg ggggcgctcac 13200
agtagcagaa gggggatcgt ggagcaggcc accctgtggt catgcactgg aagctcatta 13260
cctgacgatt tggagctcat cactgggggc ctaaggagaa tagatactga aggatgagga 13320
gtgatggcg ggggcacggg tgtctttggt ggccagaact tggggactgc tggggtgcct 13380
cactgcaggc ctctcagcg cctttatat gcttacacag gctgtttcta agagggggat 13440
acattgcata agcgttttca gactacctca tcatgggtcc ctcttttac cctctgtggc 13500
cctggtggcg cactctctgg gaagggtgcag gtggtatgcc agaccgccc tgccatccac 13560
ctgcacgtcc agagctgact tagcctcgag attgctgctg gcacctctg ccccgggaca 13620
cctcggatgt gcccgaggag atgctggctc tgtgttttct gctggagttt ggtgcgtctt 13680
ttctcctgc aagtggccac cgctcttggg tatgtcctca ggcttctgag agtcatggct 13740
gttctcagg tccttgccca gcgccaggag caaacctcc tggcactttg ttcaggggtg 13800
gatgcgccag tgttctgct gtggaccccc atctcacatg aggttcttgg gcctgcaggc 13860
tcgttcagga aacaccgct gactacgcag tgtgtgccag ctgtgtccca ggcaatggcg 13920
gggacagtgg ctgctgctgg ggttgtgtg gcttctgggg actctgggga cagctgaggt 13980
gcaaggagcc acggctcctt gaggatgcag ttggactcca ggtggaaggg atggttgggg 14040
gaggtataaa tggggtcagg gaggagacac atttgaaca atgggaacat tttaagatg 14100
ctatgtcggg aggcaacaag gtggccaacc caggtgctga ggagcccaca ccagccctgg 14160
acgtgtttg ccgctcacct ttgctgggga gtggtgggag agaggattcc gtccacgtg 14220
gtggtgtgcg cagctgggct gtgtggagct gggcgctagg aggaagggtc ttctgcggg 14280
gctagccggg ctctgccttt gaacacaatc aggtccagg ttctcagcat ccagtgcatt 14340
agaggacttc acgggcagct gtggctgctc ccttgatgaa ttgggagaag aacaaaggct 14400

tatgaaatga ggtttcatgt agatggcatt agagacgccc acaacagatt tacagagtgg 14460
agcggagacg gcggatgggt ctgggaggcc cctctgctg gccttgactg tgacagctgt 14520
cctgggaatc agcttcagg ccgccccagc agcctgactg acacacacag gggttttagc 14580
cccatcctgc gaccagctgt tgccatcatc agtgacagct gggagtggcg gtggttccag 14640
ccctgggcac cctccccacc tgctggggcc caccagggc agtcctgaca cctacaggtt 14700
gcttgagacc gcatccgagt cctgccccac cacgtgtgaa gcccagtggt tcgtgggctg 14760
aggccccctg attgcatccc cacttccctt ctgcttcaca tagctgcctc ttctaccgt 14820
ttttcagcc tcttgggcta ggaattccag tgttgtctg gctttgccc aggacacctc 14880
cttagccctc ttctgagtc tagagccccg ggggttgga gttctggccc ctgggacacc 14940
tgacgccaca ctacgttct cctgtgagcc tccagcatgt ccctcagga ccaagccctc 15000
acgttcttgc cttcccgccc acctgggctc agccagggga aggcctggct gggagcgtct 15060
ccccctgcc ctgcccttct cccctctacc ctgcccttct ctctctgcc ccgcatggc 15120
ttttatatcc tgtgccacaa gacatggctg tgtgtgaaag tggcagggtc tggcatctct 15180
gtgggtctct gaggccacg ctccagtgcc actcttcca ccgctggcc gtgccctcat 15240
gctggaggga cagcccagcc ctctccgaa cccagcccc atgtgccag ctgccccgg 15300
ccctctcccc tggaagccgg ggtcactcca gccgatgcc atggtgggga catcctgctt 15360
ccttggcctt ccagggaagg tctctttcc aatggcgac acctggctcc tgcctggagg 15420
ctggaagctg tggccctgt atgcccccc agggctctg cgctcggtg gcccgagttc 15480
ccatcaccgt catcatcacc atcatattg tcatttcgt tgtctgtgag ccggcctggt 15540
ctcccagagc agagaccctc tgaggtccag cctgagtgg ggtctccgtg ctgaccctg 15600
acggggactc aggacgtacc aggtctgggt caggagtac cccaaacct cgtgccctt 15660
gacaggcacc cctgactttt gctaagtggg tggaggtgac atcacttaca gcgggagtga 15720
tgggacaggg tctgttggt gcactgtgt cccagggatc tggggagagg ctatatccct 15780

gggctttggc actgcagagc tgtgtgtgtt tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg 15840
tgtgtgtgtg tgtgtgtgtg ttgctgtgcg cgcacatgtg tataagatct tttttatta 15900
catgaagcaa gataactgtt gctgtttcct ttgggtttt gtgttaaca gagggtggta 15960
cttcttcct cagacaacag aactctcccc tttaaacacg tgctgtcaga ggggtgggtct 16020
tgggctcatg tctgtttgca cagccgagtc agaggaaaca cagggttctt cataaaaaca 16080
ctgcacagca ggcgactgtc cagagtcagc ctgcaggacg gcagcagccc tgcccctcag 16140
agcacagcta ggggtgggctg ctttgggatc tccgtcatt cctcccagc tggcagccgg 16200
cgccggccc attccttgggt gtgtgtgtca ggggggctg cgcctgtct gtcaccctg 16260
ggaatgggac agaagctggc agctcggaga ggacagggt ggaccctgg gtggcctctg 16320
gctggacat ctattgtcc tcagacacag cctctcgggt ctagtctcat ttctgaaaa 16380
acaagtgcac agaactagag caggagtcca gagctacggc cccggggcca gatccagccc 16440
tgccacctgt ttacacacca tgctcaagct gagggtgtt tacattttt aattactga 16500
aaaaaaaaa gccaaaggag gttcatgac ccatgaaat tatatggaat tcaaaaaaaaa 16560
aaaattatat ggaattcaaa ttctagtc cataaataat ttctgagac aggtctcgc 16620
tctgtaccc aggtggagt gcagtgtat ggcatggctc gctgtaccct tgacctcca 16680
ggctcaagcg atcctcctgt ctacgctcc tgagtgtg ggactacggg tgtgtgccac 16740
caagcccggc taattttt ttaatttag taaagacagg gtcttctat gttgccagg 16800
ctttctgga actccatctt ggcctccaa agtctggga ttacaggctc gagccacgga 16860
gcccagcctg ttttgttt ttactgata aagtgtgccc ggggtgtgta gtgtgtgcct 16920
ctagcgattt gggaggctga ggtgggagga tcgttaagc ccaggagttt gaggctgggc 16980
tcaagtgatc aggaggtgaa ctatgatcat gtcattgcat tccagcctgg gtgacagagc 17040
aagaacctat ctctaaaaa tatatatata aaaagtattg ggtgtggtgg ctacgcctg 17100
tggtcccagc tacttaggca tctgaggtgg gaggatggct tgagcccagg agttgaggt 17160

tgcagcgagc caagatcgtg tcaactacact ctagcctggg tgacagagcc cagaccctgc 17220
ctctttaaaa aaaaaaacca aaaaacatgt attggaacac agccatgcct gttcagtcac 17280
gtgctctcca tgcgtcttgc tgcctcagag acccttatgg cctgaaagct gaaaatattt 17340
tctatccttt acaaaaaagt ttgctgacct ctgtcctgga aaattcatct cccaagttct 17400
cttccggcac tggcgttctt ggggtgccta aatttggccc ctgttatttc tgaactctgt 17460
tttggtctg ttcctccca ggagccagga caggcacgtt ctctgcatct tgtcccctga 17520
cgcccagagg ctggctcgg ctccagcatt ctggaaata tctggctcca ggaaaggcag 17580
aggcctcctg agtcagccca gagggaaacct gcccaggtc tgggggaggc ctgaccagc 17640
agagtggctt ttgccgatgg gttgggccgg tcaagatgtg ctgaaagtg tcctcagaag 17700
gccactttgg gattccttcc tccagtatta gagcaactga gagctgtca ttgcaagcct 17760
gatgtttcc cagttggccg ggtccaccgg gtgcctggg attctgggat ctgggtggaa 17820
agtagggggc ttgggggagt gtctgggtt ctggaatcca ggtggcaagt ggtgaggttc 17880
agggagtggc ttctgagcca ccataggggt ctctgtggga ggctctgcc atccaggaga 17940
ttccgaggc cctgccggcc cagagccagc gtcttgcgt tgcgaggct acagccagcc 18000
ccagccgggt ggaacagccc gtgcctcct ctactttgt ttggggcca cctgggagtg 18060
tgagcaagg gtagagagg aggaagtggc tgccggccgc tgccagcac cttgtttgc 18120
cttgggcct ctgtgggctc cttttattg ctctcaatg aagccaggga aatggacttc 18180
cttgctcac ttcagtcaa catgtctgga agtttggtat taaaattaag aaagtgtgga 18240
aatagagcaa gaagagaaaa atctctcaa gagataatag tgacctga gctgggcgcg 18300
gtggctcacg cctgtaaacc ccagtacttt gggaggctga ggcgggcaga tcacctgagg 18360
tcgggagttt gtgaccggcc tgaccaagat ggagaaaccc cgtctctact aaaaataaat 18420
aaataaataa ataaataaat acaaaattag ccaggcatgg tggcgctgc ctataatccc 18480
agctaaggca ggagaatcgc ttgaacctgg gaggcaaagg ttgagtgag ccaagatcac 18540

gccattgcac tctagtctgg gcaacaagag tgaaactccg tctcaaaaa aataaataa 18600
taaaaaataa aaatagtac ctctggccag gtgtggcagc tcatacccgat aatcccagca 18660
ctttggaagg aaggccgaga tgggcagatt gcttagcac aggagttga gaccagcctg 18720
gccaacatgg tggaacccca tctctacaaa aatagaataa aatttaagag gtaatagtga 18780
ccttttgta gatcgaacc tggattgctt tcttttcta aatgctgatt cttttcttg 18840
tggtgttgt gtctgtgcc gatgtccctc cccagccct gttattgtga gtggaagaag 18900
gggaaagggt tcgccgcta ctgtgagccc ctctctcac gctgggtgtc ctgggagaag 18960
cctgcacttc ttcattgtac gccagggtg ggtccctccc tggagtgggt ctgtgctgct 19020
gggatggggc caacccctca gatgtttct gagtgcaca cacagggtg tgcattcatg 19080
gccttgctg gtcttctgt tgtggaggca aaaatgtgaa gaaccctaga tgatttggg 19140
accagggtc catcacctgc tgttcattgc acaccggagc atccaggcat gggaggagag 19200
ctcagacttc caggcacggt cgcaggggct ggtctaacca tgtcccgcc cgcctgctcg 19260
tcagaaccgc ctgttgggag ctgttatcat gataccatac ctgggccctg ggctatccga 19320
ttctgactta attgctccag gttggggcca ggccgttgtt tgcgttttg ttgtttctc 19380
tgtgacgtta gccactgggc taatctgagc cctcagta cagggtgaga aactgagacc 19440
catgggggtg caaggacttg ccgaggacc agagcccctt gggggcagag ctgaggcggg 19500
gcctggcttt ggggccaga gcttcagtc ccttccccg tctcctaaca gctttttt 19560
ttgagacaag atctaccct gtcaccagg ctggagtga atggcatgat ctgggtcac 19620
tgcaatctc gtagctgcg ttccagcat tctctgcct cagcctccc agcagctggg 19680
attacagggt tgtccgcca tgccagctc gttttttt gtacttttag tagatatagg 19740
gttcacat gttggccagg ctgatctga actcctgacc tcaaatgatc cgcctgcctc 19800
ggcctccaa agtctagga ttacaggctg ggatcacact gtgcctggcc ctagcagctt 19860
tgtctgtgc catccaaca cagatgaccg aagtcttgt ttctaacaat gcattccatc 19920

tgccttacag tttgccacc tgcaaacag aggactgtc gctttctgg taagctggaa 19980
atgtaatctg gtagcaggag gcctgtggaa gcttgccctt aatggccttg tgtctcttc 20040
atcctgtcct gagagccgga gaacttggat gttgcaccta actcaacctt cctgttaaca 20100
tacagttctg caggctcatg gatcatcaga accacgtcct attcacgcg gctgtatgct 20160
tccgttggtt caggtgtttt taccttgaca gtattttctc ctcggtggct tttgcggtgg 20220
ttgcttttaa tcagcattga ctctcaaga aaaatatita gctgctacat ctcagaggag 20280
acagggtgga aagcatctga gacctgcagg ctcagactta gaaccagaag tgcctcaga 20340
gttcatccgg ccctgaccca gcgggaaatg agttcacaga gaagcgggag aactttgccc 20400
caggccctgc cgttgctcat aactgcccc ggtccttaca ttgctccag gtcctgcccc 20460
aggccctgca gttgctcata actgccccag gtccttatat ttgctccagg tctgcccc 20520
ggctctgcag ttgctctgtg tgggtgggtg gatctggagc cctccgcca ttgctgcacc 20580
tggggcaggc attgctaatt gatccagga ctcttctctg cggagcacgc cctggtctc 20640
caggcagccg ctgcctgtca gcctgcagtg gttcgggaga ggacacctgc ttgcctggtc 20700
tgttccaaat ctgcttctc atcccagcac aggtaggggg tgctatggga aagggaacct 20760
cagttggccc tgtactgct ctatcagctg gggacgtggc atcctagtga aaacatcatg 20820
gccgggcgcg gtggctcacg cctggaatcc cagcacttg ggaggctgag gaggtggat 20880
cacttgaggt cagaagttcg agaccagcct ggtcaacatg gtgaaacca tctctactaa 20940
aaatacaaaa attcgccagg tgtgtggcg ggtacctgta atccgagcta ctggggaggc 21000
tgaggcagga gaatcgcttg aacctgggag gtggagcttg cagtgagccg agatcttgcc 21060
actgcactcc agcctgggca acagagttag acgctgtctc aaaatctcaa acaacaaac 21120
aaacaaaaaa caaacaaca aagcgtcatt tatccagcac ccctggggaa catgctacc 21180
tgggttttta tggctacctg caaggtgcag gtgaagttgc tgctcttggg cattgaacct 21240
gtcttgtttg gggcagctca ggccccaggc aggttccggg ttggctctcg ttggtgtggc 21300

cctggcccat ccagacctat atttctgccg tctgcaggt gatcaatgt gatgggacga 21360
agaggcggac cctcctggag gacaagctcc cgcacattt cgggttcacg ctgctggggg 21420
acttcatcta ctggactgac tggcagcgcc gcagcatcga gcgggtgcac aaggtcaagg 21480
ccagccggga cgtcatcatt gaccagctgc ccgacctgat ggggctcaaa gctgtgaatg 21540
tggccaaggt cgtcgggtgag tccgggggggt cccaagccat ggctcagcca tgcagacttg 21600
catgaggagg aagtgcggg tccatgcctg ggcataagt ttgagctcag gtgccccgac 21660
ctggggaagg gcaggacagg aaaggtgaca gtatctggcc aaggacagat gggaaggac 21720
caaggagct gattagggag tggttatgga ctaggaatgt cgtaacaat ggtagaaag 21780
tgactaacat ttgttagca cctgctgtgt gcccgccct gccggggagc ctctgtccc 21840
acagtgacc cgtctgcaaa ttagttct tgcctactc gactgggga gcaggacga 21900
gagccgtgca tctcacaggt gccaaagctca ggactccctc ctgggtctgc ctgggtggg 21960
ctgtgttgt tggccctgtg gccacgcat gtgcacctc cacctgaaag ccaggatctt 22020
caggacgctc cccgaggagg tcgtgtctg gcacaatgat ttgtcttc ctgaaaagg 22080
gacagagta cactggagag agcagcatcc aggtgcggca gggacaggcc tggggctcgc 22140
gggcaggac tctgtctct gccggggctc cacactgcac ctgctgtca gaggcactca 22200
gtcaatctt gctgatgaag gatgagagga cagaggacgt gatgcttgc gctgcattgc 22260
ctgcagtct gggtgagatg cccgggttga ctctgtgcc cgtcgggtgg atgtgatgc 22320
agatccccg ctttaaaata cgaggagct gggaattgag ggagcaggtt ggggcagaaa 22380
gcacagcccc gtggaagcct ggagctgagg cagtgtggc gaccctgga gcagtgagt 22440
cttcttcat ggccttcac gcaccctgca gtcctcatgt aggggatgcc atccatgaat 22500
ttagtttcc cagcctcct taaaaacgcg tcatgctgg gccggggca gtgcagtggc 22560
tcacatctga aatccacca ctttgggagg ccgaggcggg tgatcatga ggtcaggaga 22620
tcgagacat cctggctaac aaggtgaaac cccgtctcta ctaaaaatac aaaaaattag 22680

ccgggtgcgg tggcgggccc ctgtagtccc agctactcgg gaggtgagg caggagaatg 22740
gcgtgaaccc gggaagcggg gcttgacgtg agccgagatt gcgccactgc agtccgcagt 22800
ccggcctggg cgacagagcg agactccgtc tcaaaaaaaaa aaaaaaaagt acaaaaaaaaa 22860
aaaaattagt ctgggtgtgg tatcacgcgc ctataatctc actactcgag aggctgaggc 22920
ggagaattgc ttgaaccag gaggtagagg ttgtagttag cccgtatcgt accactgccc 22980
tccacctggg caatagagcg agactctgtc tcaaaaagaa aaaaaaaaaa agaacattta 23040
tgccaggtgt ggtggctcat gcctgaaac ccagaacttt ggaagactga ggcaggagga 23100
tcacttgagc ccagaaattt gagagtgtct tcctgggca acatagagag acctcatctc 23160
taccagaaaa aaaaaatta gcccgcatg gtggcatatc cctgtgtcc cagctactta 23220
gggggctgac gtggcaggat cacctgagtc tggaggcaga ggtgaagt agctgagatc 23280
atgccactgc actccagcct gggtgacaga cagagaccct gtctcaaaaa aaaaaaaaaa 23340
aaaaagcatt tactatccac catggaaggt gagactgacc tgtgagtgt tgttcaaaga 23400
acaaaaata aacccagag ataagacaaa aggggtgcctc catgggggtg tgatttaaag 23460
ctgagaaatt gggtctctc cccctccct ctcacccgt ggttgctaa aggagatggg 23520
aaaaaggatt ctttttgg ctgaaatatt taacactaaa ttaaagccaa tttaacagc 23580
actttggtg atgagtgaat ttaacagact ggccaaaaat aaacgaacgg tctgtactat 23640
gtgaaaaaga ggcagctttg gccatgctgg gccaatgtga gtttcaggg ttgctgggaa 23700
tgtctgtgaa tcggaggaag ggcctagctg ggactctcag gagccaaggc cctgaggggc 23760
aactgcctg gtccctgccc tgaggcgtc actgcttct tctgggcca gatcacaggc 23820
ccggaggctg gaccactggg ctggcactct tgccgagctg ctccctgact tctgacat 23880
gctccttca gcagccttg tgacatttag ttccctgaa tgaaaaatgg ggatgagaat 23940
agctcctacc tcaaggtga atggagttag ttccgacagg tgactccctg ggaccagtgc 24000
ctggcgctg acaaggtcca gtcagagccc gcaactgctgt tactgatacc ctggctgta 24060

ccaggggaga acttggttgc cattgccagg tgttctcca ccacccccac tactgtccct 24120
gtttgatgtg tggcgggaat aaagctgtgc acattggagc ttttggcaca tcctggcttt 24180
caggtgaaag gtgcgtgtgt gttgaggggt ttagcctggc caaccagcc atgaggtcgg 24240
acctgacctg ggggtgagtc ctgagctcgg caccctgag ctgtgtggct cacggcagca 24300
ttcattgtgt ggcttggccg caccctttc cctgctgggc tgttgatgt tagactggag 24360
cctctgtgtt cgcttcagg aaccaaccg tgtcggaca ggaacggggg gtgcagccac 24420
ctgtgcttct tcacaccca cgcaaccgg tgtggctgcc ccatcggcct ggagctgtg 24480
agtgacatga agacctgcat cgtgcctgag gccttcttg tcttaccag cagagccgcc 24540
atccacagga tctccctga gaccaataac aacgacgtgg ccatcccgct cacgggcgtc 24600
aaggaggcct cagccctgga ctttgatgtg tccaacaacc acatctactg gacagacgtc 24660
agcctgaagg tagcgtgggc cagaacgtgc acacaggcag ccttatggg aaaccttgc 24720
ctctgttct gcctcaaagg cttcagacac ttttctaaa gcactatcgt atttattga 24780
acgcagtca agctaataa atatgagcaa gcctatttaa aaaaaaaaa gatgattata 24840
atgagcaagt ccggtagaca cacataaggg cttttgtgaa atgcttgtgt gaatgtgaaa 24900
tattgttgt ccgttagct tgacttcaga caccacccc actccctgt cggtgcccgt 24960
ttgctcagca gactcttct tcatttatag tgcaaatga aacatccagg acaatacag 25020
gaagactttt tttttttt ttgagacag agtcttactc tgtgcccag gctggagtac 25080
cgtagcgtga gtcagctca ctgcaacctc cgctcccag gttcaagcga ttctctgcc 25140
tcagcctcct gagtagctgg gactacagac atgcaccacc acaccagct aattttttt 25200
atatttttag tagagacagg gttcatcat gtggccagg ctggtcttga actcctgacc 25260
tcaggtgatc tgccgcctc ggctcccaa agtctgaga taacaggtgt gagccaccgt 25320
tccggcata ggaaaacttt ttgcttcta aagaagagti tagcaaacta gtctgtgggc 25380
tggccttctg attctgtaa gaaagttga ttggtggctg ggtgcggtgg ctcacacctg 25440

taatcccatc actttgggag gccgacgtgg gcataacc tgatgctggg acttcgagac 25500
cagcctcacc aacgtggaga aaccccgctc ctactaaaaa tacaaaaaaa aaattaaccg 25560
ggcatggcgg cgccctgcctg taatgcagc tactcaggag gctgaagcag gagaattgct 25620
tgaacctggg aggcggaggt tgggtgagc tgagatggca ccattgcact ccagcctggg 25680
caacaaaagt gaaactccgt ctcaaaaaa aaaaagttg attggtgtaa ccaaagcgca 25740
ttgtttatg gattgtcigt ggcagcttt ttctgccga gatgagttg gacagatctg 25800
tatgggctct aaagcctaaa acatgtgcca tccgccctt tacagaaaa gtgtgctgac 25860
ctctgttcta aagtattgga caactacaat gttgtctcat ttattattct atgatttgtt 25920
ttctgcttt tgtgttgtt gttgtgtg agatagggt tccctctgc actcaggctg 25980
gagtgcagt gtgtaatctc agctcactgc agcctcgacc tctgggctc tagtgatcct 26040
ctcatctcag cctccctagt agctgggact acaggcacac accaccactc ctggctgatt 26100
tttttttt tttttttt ttgtggagac agggttccg catgtgccc aggctggttt 26160
caaactccta ggctcaaaca cccacctcag cctcccaaag tgctgggatt acaggcgtga 26220
gccaccatgc ccagcctatt ctactgttg tattacatag cttaaaaaga tttttatga 26280
ctttaagtca caagggttct tttagaaaa aaatatatat ataggaaagt ataaaaagaa 26340
agtaaaaatt gtccataacc tctccagcca gagacgacc ttgtgacac ctcagcatat 26400
tgccttaag tctttttct ctaagatagc atttctctc atcacagtca tatgctacgc 26460
agaattctgt atcctgattt ttacttga cattacaaca ggtattgat ggcgctgtga 26520
caaactcttt ggcacaatct tttaaatgta tgaaatactc cactgcacag atgtttgctt 26580
ttaggcttaa ctgttcttt atttgcgtg tgctggttac agccgggcac agtggctcat 26640
gcctgtaac acaacacttt gagagggtga ggcaggagga tcactgagc ccagaagttt 26700
gagaccggcc tgggcaacat agtgagacc catctctaca aaaaacttt ttaataagtc 26760
gggcgtagtg gtgcatagct gtagtccag ccaccaagga ggctgagtg ggaggattgc 26820

ttgagcccca ggaggtgat gctgcagtga cctgagatta ctccactgta ctccaacctg 26880
agcgacagag caagacttgt ctggggaaaa aaaaaaaaaa aatatatata tatatatata 26940
tatatatata tatatacata cagcgacaca cacataatat aaaaatatat attataaat 27000
atataatata taatataaaa atatatatft ataaataaaa ttataaaatt atattataa 27060
gtaaatatat aatatataat ataaaaatat atattatata atatataata aaatatataa 27120
tataaaaata tatattata aataatatat aatacatact tataagtata tatttaaaat 27180
atatgtaatg tatattttt aatgtatgat atataatata cattataaaa tacacattta 27240
tattatttta tataaaatat atataaaatc tccaagttgc ttttcctaaa aaggtgtctt 27300
gctgcatttc aaacattcat ttaaaaactt gaatgctggt gatctggtcc agaattgtt 27360
cagtagctgc tgccagtggc caagcatctc gggagatgtc taaaaaacac gctggttctg 27420
gcctggcgtg gtggctcacg cctgtaatct cagcacittg ggaggctgag gcaggtggat 27480
caactgaggt ctggatttcg agaccagcct tgccagcttg gtgaaacccc atctctacta 27540
agaatacaaaa aaaattagcc aggcgtggtg gcattgtcct gtaatccac ctacttgga 27600
ggctaaggct ggagaatcgc ttgaaccag ggggcagagg ttgcagtgag ccgagatcgc 27660
accattgcac tccaggctgg gcaagaagag cgaaactccg tctcaaaaaa aaaaaaaaaag 27720
atgctggttc ctaaaatgtg gcccttttc tctcacctg ctgccagacc atcagccgcg 27780
ccttcattgaa cgggagctcg gtggagcacg tgggtggagt tggccttgac taccctgagg 27840
gcatggccgt tgactggatg ggcaagaacc tctactgggc cgacttggg accaacagaa 27900
tcgaagtggc gcggctggac gggcagttcc ggcaagtcct cgtgtggagg gacttgaca 27960
acccgaggtc gctggccctg gatccacca aggggtaagt gttgcctgt cccgtgcgtc 28020
cttgtgtica cctcgtatga gacagtgcgg gggtgccaac tgggcaaggt ggcaggctgt 28080
ccgtgtggcc ctactgatt agagctgtac tgatgtcatt agccttgatg gtggccagga 28140
ctggtagggc cctcagaggt catggagtgc ctctgtggag cgggtgctga ggctgtatca 28200

ggcacagtgc tggctgcttt cacctgggcc gtctcaccga agtgtccatg gagcctgcgt 28260
agggtgggta tctgtgtcga tttacagat gcagaaacag gctcagagaa accgagtgac 28320
ttccctaagg tcacataccc agttagagca gagctgggcc aggaagtgt gtctcaggct 28380
cctgaccagg tctccttgc tgcactctt gccaaaacca tgatccagaa ctgacttga 28440
ggccccgga cctcaggctc ctccgaaatg gcctcttga ggctgctgag ccacagctta 28500
ggaccacct cgagaggcaa atgtgcttgg agctgccagg cgtcctgggg gcctgcctt 28560
gggcacgggg ttacagacagg cccagatgt gtggggcgtc ttctggact tgagtttct 28620
ttctgtgtg gtggacacag tgctacccc ttaaagcacc tgtgatgtgt gcagagccc 28680
aatccctgcc tgtgcctgt tctgctaggg aaggaaggaa gacttcagga tggcaggaca 28740
acagaaagag gtccaggtt tagagcaagg gcaggtcaaa cttagaaaat tctggaatga 28800
ggatgtgcat ttctcttct ggatctgcta aaagaagagg gaaggagggg ctgctggggg 28860
aggagcccag agccgagttt acatccgat cccgcaaggc ctcccctgcc ctgaggtctt 28920
gtttgtgat gtgcttgtgt ccatcctggt ttctgccgtg tcccaacat ccggccaagc 28980
ttaggtggat gttccagcac aactcacc tgtctgtgca cctgttttg tgtccgtaag 29040
tgggtattta ctacacttac gtagtagcca ctgtgggaat tcaggagagt ggcgcagtga 29100
ccaccctgg agggatatgt gtgtggcagg ggtcgagggt ctgcccttc cctgttcct 29160
gcgcgtggct ttctccagga cggggagggc tgagctgaag aggtggggac agttgcgtcc 29220
ccccgccacc cactgtcctg cggtagagc agactcactg agcctgccct tctccctgt 29280
gccttcacg tacatctact ggaccgagt gggcggaag ccgaggatcg tgcgggcctt 29340
catggacggg accaactgca tgacgtggt ggacaagggt ggccgggcca acgacctcac 29400
cattgactac gctgaccagc gcctctact gaccgacctg gacaccaaca tgatcgagtc 29460
gtccaacatg ctgggtgagg gccgggctgg ggccttctgg tcatggaggg cggggcagcc 29520
gggcgttggc cacctccag cctcgccgca cgtaccctgt ggcctgcaag tccccaacc 29580

tggcaggagc tgtggccaca cccacgactg cccagcagcc tcacctctg ctgtgggagt 29640
tgtccccgtc caccctggg tgcctttgct gcagttatgt cgggagaggc tctggtgaca 29700
gctgtttcct gtgcacctgc tgggcactag gtcccagcta atccctgtgc caggactcta 29760
atttaccct aacacacatg gtggttttca ttgctgggga agctgaggcc tgagcacatg 29820
acttgcctta ggtcacatag ctggtgagtt caggatcccc cagagatacc agggccagca 29880
ctcgatcccc acccagccct gaacccacc atgtgctggg attgtgctgg gagtgtccac 29940
acgcctggga cccagggt ggtgctctca tctcctttt ccagatcatg agaatgaggc 30000
tcagggaagt ttgaaaaaaaa cctatcccaa gtcacacagc aacaggagca ggatttgaac 30060
ccagaaaagg ggaccgcaca ctctgttctg ctagagtagt tagctgtcct gggtgatatg 30120
gcaggtgaca ggggcaactg tgcttaacaa aggaaccccc atccccctg ccaagttggg 30180
agactagaag gtcaggggca gaagctctga agggccagggt gcagtggctg acacctctaa 30240
tcccagcact ttgtgaggcc aaggcgggca gatgatttga gcccaggagt tcaagatcag 30300
cctgggtaat gtagtgagac gccatctcta caaaaaaatt tttaaaaat tagctgggca 30360
tggtggttca tgcctgtagt ccaagctact tgggaggctc aggtgggagg attgcttgag 30420
cccaggaggt tgaggttggt gtgagctgtg atcatgccac tgcactccag cctgggcaat 30480
agagtgagac cgtctccaaa aaaaaaaaaa gaagaagaaa aagaagctct gaggctccaa 30540
gtccccaggc accccttggc ttgagggcag acaaggagg agagggtcac ctgggcagcc 30600
ctgacttttg tcccctggca aaggacctt cagtgcctt ggccctagga gagcctctga 30660
gcacgtcagc catgtcgaac cgctcaggaa gggcagcaag aatttgctt ctgacctctg 30720
cctctctac tcgcatctg cactgggtgt ggttgtgcc attttacaga tgaggaggct 30780
ggggcatcga ccagctgaat gccttgccc aggtactgcg taggcagagc tggcagttga 30840
acccgtgtc ctggttgctg ctgggggtgg gctgcacct gactgtgag gccagtagca 30900
aggttgcac gtgacttctg gaccgtcacc cagctctgca gcacatccc tgaccagct 30960

tggcaggagc tgtggccaca cccacgactg cccagcagcc tcacctctg ctgtgggagt 29640
tgtccccgtc caccctggg tgccttctg gcatgtatgt cgggagaggc tctggtgaca 29700
gtgttttct gtgcacctgc tgggcactag gtcccagcta atccctgtgc caggactcta 29760
attcacct aacacacatg gtggttttca ttgctgggga agctgaggcc tgagcacatg 29820
acttgctta ggtcacatag ctggtgagtt caggatcccc cagagatacc agggccagca 29880
ctgatcccc acccagccct gaaccccacc atgtgctggg attgtgctgg gagtgtccac 29940
acgcctggga cccagggct ggtgctctca tctcctttt ccagatcatg agaatgaggc 30000
tcagggaagt ttgaaaaaaa cctatcccaa gtcacacagc aacaggagca ggattgaac 30060
ccagaaaagg ggaccgcaca ctctgttctg ctagagtagt tagctgtcct gggatgatg 30120
gcaggtgaca ggggcaactg tgcttaacaa aggaaccccc atccccctg ccaagttggg 30180
agactagaag gtcaggggca gaagctctga agggccaggt gcagtggctg acaccttaa 30240
tcccagcact ttgtgaggcc aaggcgggca gatgattga gccaggagt tcaagatcag 30300
cctgggtaat gtagtgagac gccatctcta caaaaaaatt tttaaaaat tagctgggca 30360
tggtggttca tgcctgtagt ccaagctact tgggaggctc aggtgggagg attgctgag 30420
cccaggaggt tgaggtgtg gtgagctgtg atcatgccac tgcactccag cctgggcaat 30480
agagtgagac cgtctccaaa aaaaaaaaaa gaagaagaaa aagaagctct gaggtccaa 30540
gtccccaggc acccctggc ttgagggcag acaaggagg agagggtcac ctgggcagcc 30600
ctgacttttg tcccctggca aagggacctt cagtgcctt ggccctagga gacacctga 30660
gcacgtcagc catgtcgaa cgctcaggaa gggcagcaag aatttgctt ctgacctctg 30720
cctctctac tcgcatctg cactgggtgt ggttggtccc attttacaga tgaggaggct 30780
ggggcatcga ccagctgaat gccttgctcc aggtactgcg taggcagagc tggcagttga 30840
accctgtgtc ctggtgtgtg ctgggggtgg gctgcacct gactgtgag gccagtagca 30900
aggtttgac gtgacttctg gaccgtcacc cagctctgca gcacatcccg tgaccagct 30960

catccaggcc gcatgcaaac ctgttgccag gcgagaaacc agtcaccgca cagctgtggt 31020
tgctgaaat gattaagctc attaatcacc ccggagttag gacagactca gatgaaaacc 31080
agcaaaagcc ctggaaactc atgtgaccct gccaatgagg gcggccatgt gcattgcagc 31140
ctggccgtca ctctcggta cgtgttttgg acttaaacgc tccggtatgt tactgagtgc 31200
ttgattaata acatggaagg cctgggtctca ttgctgtggg agtgaaggat gcacagccag 31260
gcctgacatg atgagaacaa gaacctggag tctcgtgcc tgggtggtta tcttggccct 31320
gccacttagc aactgtgtga ctgtagccag gtcacttaat ttgctagat cctgcctgcg 31380
cttcagtga tcttctggtt ttccaagggt ggccaaacac ttaaggcat tcatgtggtc 31440
gctaggctgc agggttgaac cctgggtcac cccgcagggc gccgtgtgct ctgtggcctg 31500
gctgtgcctt tgctgacacc gtgcccgtgt gtgttcacgc aggtcaggag cgggtcgtga 31560
ttgccgacga tctccgcac ccgttcggtc tgacgcagta cagcgattat atctactgga 31620
cagactggaa tctgcacagc attgagcggg ccgacaagac tagcggccgg aaccgcaccc 31680
tcaccaggg ccacctggac ttcgtgatgg acatcctggt gtccactcc tcccgccagg 31740
atggcctcaa tgactgtatg cacaacaacg ggacgtgtgg gcagctgtgc cttgccatcc 31800
ccggcggcca ccgtgcggc tgcgcctcac actacaccct ggacccagc agccgcaact 31860
gcagccgtaa gtgcctcatg gtccccgcga cctcactccc tcgttagatc aggtgtgtc 31920
tgggagctga cgctgaaagg agcttctcat ctgggggtcc tgggtgtaca tagatggtg 31980
ggtaggttgt gcactgcaca agctgcatga tgctacctgg ggggtccagggt ccaggctgga 32040
tggacttgtt gcttcacag gacatagata aatggccaaa actcctcagc tgggaaggctc 32100
tgggcaggat ctttgggtgt gaaaaccagt cacaggggaa ggggtgctgc tcatactgcc 32160
agcacagtgc tgagtgttt ccatagcgt cgttactcc tcaagcctgg aggggtggga 32220
gtagcatggt cccatttcac gtacaaggaa ccgcatgcac agagaggtgt ggcaacccat 32280
ccaaggccat acaactgggg tgggttgagc cgggggtgac tgtggcaggc tggctcaaga 32340

gtccctgctc ctgaaccctt gccaggcagc ctggcatcag ctcggggaat tttgccctg 32400
acccttggaa gcaagtgggc ctcttgttc tcatgtcagt gatgagaaga gtgactttcc 32460
tatggcccct ctggagtaca ggtgtttcct gttggcgggc tcttccccca tgacatcagc 32520
agcgagctgg ttatgattcc ctacgcagaa ctgatagtt tataaagctc tttgtcatcc 32580
aggccccgtt ggagtctcac gcagacctgg tcgcaggcgg ggctggtctt gcctgtccca 32640
gctgcatgga tggggaactt gaggcttga aaggttaagg ggctgttga ggcccaggct 32700
ggcaggagat gggcctgggc cagagtctgg gacttccat gcctgggctg tctttggtcc 32760
tgttctcac catccctccc tggggccatg acctagaga gccaaatga ggtgcaggta 32820
accacggca aggaggggtt gccatgactc agagtcccg tcctgtggcc ggcagtacct 32880
ggtgaacga ctggatttc agaccagcca ctgtagcccg ctgacggtgc gctcgaagt 32940
ccacagcttc tgaagccagg caggactcag gccaggagac tctgttagct gttgagaggg 33000
agaggccaac ggatgttctg gttctgtag agagctgggt cticggatcc tggtagcagt 33060
gcactgagag gaggcccagc ttgattctgg ggctgccttg tggtagcatg tgctgctcac 33120
tgacaccctc gaggagtgtc ttctctggg ctgttgact gtgcccgtt ttccgcagt 33180
cactggtgca cacataggca catagcaaac cgcacacaca gtcgtgggta tgagtttcac 33240
tacattccac caccagtgtt cactaccatt acctgccttc cgtcttaagt gttcatcatt 33300
taaaaataaa ttattgggc tggacgcgtt ggctcatgac tgttatcca gcactttggg 33360
aggctgaggc gggcagatca cctgaggtca ggagtcaag accagcctgg ccaatatgtt 33420
gaaactccat ctctactaaa aatacaaaat tagctgggca tggtagggca tgcctataat 33480
cccagctact caggaggctg aggcaggaga atggcgtgaa cccgagaggc agagcttaca 33540
gtgagcccag atagcaccac tgcagtccag cgtgggcaac agtgcgagac tccatctcaa 33600
aaaaaaata aataaataaa agaaaaataa atttatgatc tatttcaaaa ataacacatg 33660
tactttgaaa cagcagagac acatatgaca cggagaatga aattcccat agcgcacccc 33720

caagagacag ccctggtccc cccgtctttc ccgtggacct ccagcggggc agatgctgag 33780
ccgcctgttg tcgagtggcg tgctatcccg tctccagct cctctgtggc ttacagacac 33840
ccacctgcag ccctgtcttt gcctcctcta ggcgccacca ccttcttgct gttcagccag 33900
aaatctgcca tcagtcggat gatcccgac gaccagcaca gcccgatct catcctgccc 33960
ctgcatggac tgaggaacgt caaagccatc gactatgacc cactggacaa gtcatctac 34020
tgggtggatg ggcgccagaa catcaagcga gccaaaggacg acgggaccca ggcaggtgcc 34080
ctgtgggaag ggtgcgggggt gtgcttccca aggcgtcct ctgctggtt tccaggctgc 34140
tggccctgtc cttagcagag ggaggaaaca gaggatggct ctgggtgaat gatgactgg 34200
gcttcgatta tgtagtcaca gggtagtacc ctgagatgcg tggaacccc agactgtgat 34260
tatatgtaga aactgggttt ccccgttgtt taagtagtca tgggtgggtc agaccccaca 34320
ggacttttgt ctttcaaga aagaaaatgg tcgtgtgtca tgcaggggta gttggtactg 34380
gttaatccag gtttatcctt tattttgtgg gaactgtaca gtcatttctg ctacaatgct 34440
gtatatgctc ttctgaaaga cacctatgca aaatgcaca gtaaaaatga cacaactcat 34500
agggaaagcg gggccaggcg acagccctca aaatctccat caatgacatg taagaaaaga 34560
gaggaacctg ggaaatagca aagtgccttt tgcacattaa atggttagct atatcccaca 34620
atactgtgca ttctgaaacg ttaatgtgc aataaatacg gcacttcacc ttgggaagat 34680
ctggagtgg cttatgagtg tggaagggtg tagcgcatga gttttgtga aacactggaa 34740
ggaggattgt gggaaatcaa atggaaagt ctcaccccag gcgtggagaa gagtgggtca 34800
tggccccagc agtgagcca gggaggtcag agacggaggt gtgtgtgtgg gtgtgaccct 34860
gcgcagtcc ctgccggtg tagtttttg cattcgctta atgtttctg tggaggaaat 34920
tgtgcatgag caaatgtgaa accgtgctgt gctcaaattg tcctaataca tcattgcatt 34980
ggaacagatt ggctttttt tttttttt tttttttt ttttgagat ggagtctcac 35040
tctgtcacca gcctggagtg cagtggcatg atctggctc actgcaacct ttgcctccta 35100

tggtcaagtg atttctctgc ctacgctcc tgagtaactg ggattacagg catgagccac 35160
cgcgcccggc cagatttgca ttttgaaac aactgctagg ctgggcgcgg tggctcacac 35220
ctgtaatccc agcactgtgg gaggccgagg cagggtggtc acctgaggtc aggggttcga 35280
gaccagcctg gccaacatgg tgaacccccg tctctactga atatacaaaa atcagctggg 35340
tgtggtggcg ggtgcctgta atcccagcta ctgaggaggc tgaggcagga gaattgcttg 35400
aaccagggag gcagagggtg cgggtgagccg agatcacacc attgcactcc agcctgggca 35460
acaagagcaa aactccatct caaaaaataa aaaatagaaa aacaagtgtg gtagcggaag 35520
tgagcacttt gcggagtcag gcttgtgtgg cctgttcac aaatgatgtg ctcacggtgg 35580
ctcaggccc acctggagtc tgcagcatgg ggcacaacag gttcattagt gtagaattcc 35640
aggacaggcc tggctcctaa gcagccttct ttacaaaaa ctgcagagcc cgcctgtatc 35700
ctagcacttt gggaggccga agtgggtgga tcacgaggtc aggagtcaa gaccagcctg 35760
gccaacatgg tgaacccccg tctctactaa atatacgaag attagctggg tgtggtggca 35820
cgcgctgta gtcccagcta ctgggaggc tgaggcagaa ttgctgaac ctgggaggtg 35880
gaggttgag ggatctgaga ccatgtcatt gcactccagc ctgggcaaca gagcgagacg 35940
ccatctcaaa aaaaaaaaaac ctacagagcc acacggcctc ttctccacc gagtgttgg 36000
gtgggagctt gtgttattgt ggtgaaatct tggactttc ttgaggcaga gagaggctga 36060
ggcctggag agactttcac atgggtcgcc atgtccgccc tcggttcgc tgtgtgctc 36120
cccatctgaa ggctggtgcc gtccagacag gctggacgcc ctttccacc agatccttc 36180
tccgcagca gtttctagt acgtgtact gtgaggtctg tgccttggg tgatggcaaa 36240
agtcagccga atgaaattc agagccatgc ctggctccct ggagcttctc tctgggcag 36300
ctgtgatcat tgctctgct gtggtgtggg tggtggaat ggattcctt catcttgctt 36360
gctacagggt actgtcacgt ggagtcctt ggagagagg acgtgttaat tgatggatgt 36420
ggctcccatg ctgagaaagc tctgggcgt acattgcctt agagtttcat tggagctgcg 36480

ttcttttatg gtgtctgcta ggcagaagtg atgaagactt ggaagaaaac ccagaagggt 36540
ttccacttaa ttggaaaat gtgctttcc cctcctgtgt ctttgctaa ggtccagcct 36600
cctgcagcct ccccgctctg tggactctgg ctttgattct ttattaggag tccccctgct 36660
cccccaaaag atgggtgcta aattatcatc caattggccg aggttttgtt ttctattaat 36720
tgtttttatt ttttattgtg gtaaattat ataacataaa atttgccatt ttaattgttt 36780
tgttattgtt gttttgaga cagggtctca cccagtgcc caggctggag tgcagtgggtg 36840
cgatcatggc tcactgcagc ctacgcctcc agggctccag tgatcctctc acctcagcct 36900
ctctagtagc cgggactaca ggcatacact accacatctg gctgattttt tgtatttttt 36960
ttttattgta gagaccgct atgttgcca ggctggtctc aactcctgga ctcaagccat 37020
cctccacact caccctcca aagtgtggg attacaggca tgagccacaa caccagcca 37080
ttttaatttt tttttttt ttgagatgg agtctcactc tctgcccag gctggagtgc 37140
agtggcgtgg tatcaactca ctgaaccctc tgcctcccag gttcaagcga ctctcctgcc 37200
tcagcctcct cccgagtagc tgggattaca ggtgcccatc actatgcctg gctaattttt 37260
gtatttttta gcagagacgg gggttcacca tgttgccag gctggtcttg aactcctaac 37320
ctggtgatcc gcccgcctcg gcctccaaa atgctgagat tacagggtgtg agccaccgtg 37380
cccggccttt tttgtttt gagacagggt ctggcctgt caccagact ggagtgaat 37440
ggtgggctct tggctcactg cagcctccgc ctcccaggt caagttgtgc acctccacac 37500
ctggctaact gtattttatg tagagacaga ttaccatg ttgccaggc tgggcttgaa 37560
atggactcaa gcagtccacc cacctcagcc tccaaagt ctgagattac aggcgcgagc 37620
caccgcacc agcccatttt acctattctg cagttgacag ttacgtggca ttacgtcagt 37680
tcacgaggta accatcactg ccattcatct ccagactact tcacctctc ggcagatgtc 37740
cgaaactgtc cgcattgaac acactcctca tctcctctg acagccacca ttctacttig 37800
tatctctctc tgccttctct aggtacctca tgtaagtga attatacaa tatttgcct 37860

tgtgtgactg gcttctttca tgtgacatgg tgcctcaag gttcatctgt gttatagcct 37920
gtgtcagaat ttcttccctt aaagcctgaa taataacccg ttgtaaaggc tgggcgcggg 37980
ggctcacacc ctctaattccc agcattttgg gagtccgagg tgggcagatc acttgaggtc 38040
aggagtttga gaccagcctg gccaacatag tgaaaccctg gctctactaa aagtacaaaa 38100
ttagctgggt gtggtggcgc gcacctgtaa tcccagttac tcaggaggct gaggcaggag 38160
aatcgttgt acccgggagg cagaggttgc agtgaaccaa gattgtgcct ctgcagtcca 38220
gcctgggtaa cagagtgaga ctctctgtct caaaaaaaaa aaaaatcatc ggatggatgg 38280
acggaccact tcttgttatt tatccatcca cgggtgctag gtttctcca ccttggttg 38340
tcgtgaataa ggccactatg aacatttct tccgtgtga aggttttga ctagtgagga 38400
aaaggcgtgt ttgtggtgtt gcataggatt ctggaagaa agtttgcact aaccataagt 38460
attgtacta cattaaaatg aaagctcagg ggccgggcgc ggtggctcac gcctgtaac 38520
ccagcacttt gggaggccag ggcgggcgga tcatgaggtc aggagatcaa gaccatcctg 38580
gccaacatgg tgaaaccccg tctctactaa aaatacaaaa aaactagcca ggtgtggtgg 38640
cgggcacctg tagtcccagc tacttgggag gctgaggcag gagaatggcg tgaacccggg 38700
aggcggagct tgcggtgagc cgagatcgt tcactgcact cgagcctggg caacagagca 38760
agactccgtc tcacgcaaaa ctctgtctca cgcaagactc cgtctcaaaa aaaaaagag 38820
ttcagggttt atgaaactgg ccagccgcgt aaagtttgc gtgtgtttt tgtgcccggg 38880
aggagtgtgg ccagggtgtc acgtcacaca gtacacgtt ctcagatggt ggttctccag 38940
actgctgtcc caaagtctgt tttgcatct ggtcccaca gaccaccct ccacggtgag 39000
cctgattttg gccagggtag ctggaatctt gctgtctt cagcccggca gctgtaccag 39060
tccagggtcc acagctagt gcttttagga aggaattgt tcagttggct ttgacacatg 39120
gcccctagg gtccacagct ctgtagtgt gtggatgtt ttatctacaa agacacatga 39180
tccttcgtgt ccagatgaaa gtgatgatgt cttgcagct gccagcaag gctgtgtgtg 39240

tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgggtgtgtg gtggtgtgtg tgtgtgtatg 39300
ggggagggag gcaccccttc catctggggg tgtgtgtgtg tggggtgtgt gtgtgtgtgt 39360
gcgcgtgtgt gtggtgtgtg gtgtgtgtgt gtgtatgggg gaggcaccct ttccatctgg 39420
gtccaagaga ctgggcctgg ggaagacgct tctttttatc tacttagaga ctttgtttta 39480
tttgtatttt tttagacag ggtctcactc tgcacccag gctgggggtat ggtgatatga 39540
gcatagctca ctgcagcctc ggccctccag gctgaagcga tctcccacc tcagccttct 39600
gaatagctgg gactgtaggc gtgcgtcacc atactgagct attgttttt ttgtttggtt 39660
ggtttaattt ttttgatac agatggagtc ttgctatgtt gccagacta gtctcaaact 39720
cctgaactca agtgattctc ccactcagt ttcccgacat tctgggatca caggtgtgag 39780
ccactgctgt ctccctgttt tattaactgc tgaagacct agataaagaa agtctgaaaa 39840
gacttactat cagagcacca tcctaagatg attccctctg actcaatgga gagggagggg 39900
agcttttctt tcaggcctgg gtggcaggag cccagggtgt ccaggcccca ttgccccag 39960
gccaaatcac tcgggaactt ggatgcagct gtctttcagg gtaacccaaa ggaaccagat 40020
ccccgcaggc agtaggcttc tgggctgtcc tctctccta cgtcagctca gtaagagccc 40080
ttcgaaggga tgctgtgtcg gaggccccaa aagcccaggc tcatccctga gatgcacagg 40140
gtgggctggg cttaggcagc gctcgagcat ctctggacg gtgacccag agagtgtgga 40200
gacggagagt ccttgagagt cactgagaga cgtggctgcc ctgccttccc aagaggggct 40260
ctgagtcatt cccacactc acctgcccct acccaccctc acctggcccc cagcctcacc 40320
taccceccaca tctgtaccga tccctttacc cgcaccttc ctaccaccc tcacctcccc 40380
tgtacctca cctccccac tcaccgccc ctgcaccctc acctgtccc caccttacc 40440
taacccccac cctcacctgc cctccctca cctggcctcc ttccgttggg gaaggggttg 40500
taaggggagg ccccaaact gtctgtcctg gtgccctga gagaaaacag tacgtgaggg 40560
ccgcagtcca aaagcttgag tcctggaagg tggaggagac agggatgtgt tgggaagggc 40620

cccatggtct tggatccctt ctcgactgtc aatggggcct tcatgggagc gccagtctag 40680
tgatgcacag ctgggtgccc ggcgggtggc tgaggaggcc taaagtccga ggcggcaaga 40740
gtcttccag aggctgtgt cctaactgct ctggcatact caggcgggca cgtagttagg 40800
agctgattgg agaggagaga cccccacacc aatactggga ttgactttc aggctaaact 40860
tgagaagtgt ggcctctgct gtcctgccag agctctccag ccagtgccca gggctctcca 40920
gccagtgcc ggggtctcc accagtgcc ggggtctcc gccagtcca ggggtctccg 40980
ccagtgccca ggggtctccg ccagtgtca ggagtcttg tttcttgc ttacagccct 41040
ttgtttgac ctctctgagc caaggccaaa acccagacag gcagccccac gacctagca 41100
tcgacatcta cagccggaca ctgtctgga cgtgcgaggc caccaatacc atcaacgtcc 41160
acaggctgag cggggaagcc atgggggtgg tgctgcgtgg ggaccgcgac aagcccaggg 41220
ccatcgtcgt caacgcggag cgagggtagg aggccaacgg gtgggtgggg gtgctgccc 41280
tccaggcgtg cccgccgtgt cttatgccga atgccagcct ctcacaggct ggggagactt 41340
tccacctggg gatccaatgg gtggcttcc aggtcccaa aagcaaacac aggttttca 41400
cagcccgtcc gggaaagcag aaagcccaa ggggttgaa ggggaaagg ggagctctgc 41460
tgagaggta caaggcagcg ctggccgacg ggagttgcag ttgataggtt ttgatcatc 41520
cttgtaaac ttgaacctg tgcagaaatc cttccacgg catgggggt gcctgttgac 41580
tcgtcctgt tccaccacag ggagctcctg ggcttctcc tcccagaggc cccgacgct 41640
cccacctgtt ggtcgtcaga gcttctggtt ggtgggaagg caccaggac ctgaggtct 41700
ccagagagaa aagccaggga aagagggaga ccgaaacca tgtgacatga aactcaggct 41760
ccaaactgag cacgggaacg ttggggaca ggagcgcgat ggccttctc agatagctgg 41820
ggggctggca tgaagacggg agctacagcc agcacaggtc ctgggccggg agcccagaga 41880
ttgagccctg actctgtcac ttactggcca cgtgacctg ggcgggtggc atagcctct 41940
ggagactcag ttctcatt ggtaggagt acggccacag tgggtcgcc tctgcagcac 42000

acggggggct cgggtgggcgg aagccccggg tctataaggc ggctgtgcag gagccagccg 42060
agctggtctc ccaacagcca ggggtccggg gtccttagca gctgtggggg gcctgcacct 42120
gtttcccatg gctgctgtca gaaattacca gaagccaggt ggctgagagt aatggacact 42180
tgttctctca cagttcctga ggggtgaagc ccgagatcga ggtgtgggca gggccctgcg 42240
ccctctgaag gctctgaggg aacctttggg cttctggtgg ctccaggcac cccttgactt 42300
gtggctctgt cactccagtc tctctgtctg gctgcacatg gcgtggcctc ttctgtacca 42360
ttgaaggaca cttcagttgg atttagggcc taccctcacc cattgtggtc gtatcttgat 42420
ccttcatgac atttgtaaag accctgcttc caaataagct cacattctga ggttctgggg 42480
tgagcgggaa ttggagagc attgtcaac tagtatagaa tgtgacctgt cagcctcggg 42540
cagccctgag aggcaggggc ttccacagc ccagctgggt gccctgggct ccgtgctgtc 42600
cgaggagacg ccatccccac acccgtcctt caccgccac cctcccag gtacctgtac 42660
ttaccaaca tgcaggaccg ggcagccaag atcgaacgcg cagccctgga cggcaccgag 42720
cgcgaggtcc tcttaccac cggcctcctc cgccctgtgg ccctggtggt agacaacaca 42780
ctgggcaagc tgttctgggt ggacgcggac ctgaagcgca ttgagagctg tgacctgtca 42840
ggtagcgcc ccggggcctg ccctaaccgc agacaccg cttcattgt cagtaatggc 42900
agcagctgcc acattgtccg agacctgccg tgagcccagt gccgcgccag gggctttgtg 42960
tgtagcgtgt ttgtctca cactgacagc ttaggctgg ggttctgagt gagccccaca 43020
gggcagaggc agaaaatgag tctcagagag ggtgagcgag ctgcttgggg cccacagca 43080
ggagatggag caggactgca gcctagcctc tgccccagc acctgcgcaa gaagctgctc 43140
tgctctggac tgtgttaggc tgcgaggggt ggagagaaat gagagtgggt gcttagagag 43200
ggggcgcagg tccccatggc ttctctct atgatgaggt agatgggtga agggaggggc 43260
catgcttgca ggggccagt accgagggcc gccgttgga ctgatggcct tcatcccag 43320
cccagcccag gtgggagcag ggctttccga gggcttgtct tgggtcggcc tgcttcagg 43380

gactctgctg cagctccac cctgtccaa agcatggaat ccccaggct ccctggcagt 43440
cctgtcaacc tctgtctcc caagctgagt gtggggcaag ttctggaggt cagcactgct 43500
cagggggggcc cacgggctgc ttgcaggggc caaccgctg accctggagg acgccaacat 43560
cgtgcagcct ctgggctga ccatccttg caagcatctc tactggatcg accgccagca 43620
gcagatgac gagcgtgtgg agaagaccac cggggacaag cggactcgca tccagggccg 43680
tgtcgccac ctactggca tccatgcagt ggaggaagtc agcctggagg agttctgtac 43740
gtgggggctg gcagtggggt gggcaggggt gcctctaaac ccgaccctg gaggaggctg 43800
gaggccagtg caagatcctg tgtggcctca gccaggcgtt ggtctctgcc agatgccaac 43860
tgttggccgc tgggggtcag cgacatgtcc gaatgtccg aggcctctga ggttgtttc 43920
ttttgccga gaacaaatca ccacgaacag cgttttaaga caacaccaac tctttttt 43980
tttttttt tgagtcagga tctgtctg ttgccaggc tgggggtgcc tggtgcaaac 44040
acagttcact gcagcctcga cctctgggct taattaagt aacacctgc ctcagcctcc 44100
caggtagctg ggactacagg tgggcaccac cacacctgc taattttt ttgtagagac 44160
gggggtttccc catgttgccc aggtcgtct gcaactcctg ggcacaagct atctgcctgc 44220
tgtggcctcc caaagtgcta ggattatagg tgtgagccac tggcctgaca acaccacgg 44280
attgtctctc agttctgtaa ggcaaagtcc aggcacagcg tggctcacct gggttctctg 44340
ctcagggtct cacggggcca gaatcaaggt gtcaggaacg ctgggccctc agcggaggct 44400
ctgtggagaa attagcttcc ttgtcactc agcaggtagc agttgtggga tcgaggttct 44460
gttttctctc tggttattgg tcggggacca ctctcagctc ctagaggcca ccacaggtcc 44520
ttgccccgtg gccctctctg cctcagcagt gggggctccc tgcgtcagtc cctccacac 44580
cttgagtctc tctgatttgc ttctaaagg ccctgtgatt cggctcagcc accttagat 44640
taggttagcc tccccttga tagactcaa gtcggctgat taataacctt aatcacatct 44700
gcagaatccc ttctgccaca taaggtcatg acgccgtgct ggggactggg gtgggaaatt 44760

acgggggtcat ttaggattct gcctgccact gccttgctgt gtcccagggc ttgggggagg 44820
ggcctccaca gctgggacca cagtccttcc tcccctccat ggtaaccatc tgaggattac 44880
ttgagaccag cctgggcaac atggtgagaa cccatcccta caaaaaatac aaacaaaaag 44940
ggaccaggct gggcttggtg gctcatgcct ataatcccag cactttggga gaccaagggtg 45000
ggctgatcac ttgaggtttg gagttcgaga ccagcctgcc caacatagt aaatcccgtc 45060
tctactaaaa atacaaaaat tagctgggtg tggtgccagg cgctgtatt cccagctact 45120
ggggaggctg aggtgggaga attactgaa cctgggaggc ggaagttgca gtgagccaaa 45180
attacgccac tgcactccag cctaggcaat agagtgagac tccgtctcaa aaaaaaaaaa 45240
gggccagggg tggtagtgac aaagagacc tatcccaaaa aaaccgaaca ctgaatcctt 45300
gagactgagt aaggacactg tgaattttt ctgggtgggg cagggaacag agcgtcttct 45360
gtcatttctt ccacctgggt gtggtcagct ctcccctcaa gctgcctcct ctcttctca 45420
ttgtccgggt gtiggacaca ttgggtaac tggatagaat aacgcgagt cccagggact 45480
tggtccattt gctattttat tttattttat tttattttat tttattttat tatttttta 45540
tttattttat tattgagatg gagtttcgtt ttgtgcgcc aggctggagt gcagtggcgc 45600
gatctcgggt cactgcaacc tctgcctccc aggttcaagt gatttccta cctcagcctt 45660
ccaagtaact gggattacag gcacccacca ccataccagg ctaattttt tgtattttta 45720
gtagagacgg gtttcgcca tttgcccag gctggtcttc aactcctagc ctcaggtgat 45780
ccacgcacct cggcctccca aagtgtggg attacaggca tgagccacca cgcctggcac 45840
catttgctat ttaattccc atgtgtatta gtgtcccag gctgctgtaa caaatgacca 45900
caaactggat ggcttaaagc aacagaaatg gattcccca atgtgtgga gaccagaagc 45960
ctgcgaccaa actgttgga gggctgtgct tctctgggg gctccaggga ggatctattt 46020
gttggccctt ccagtgtgt ggggtccagc gtccacact tgtgatgcg ccgcctcaac 46080
ctctgcccac ctcatgtgt ccatctcctt tgtgtctgcg tctttacctc ttctcttgt 46140

ctgtgttgcc tcttataagg acgtttgtca ttgggttag ggcccaccca aatcatccga 46200
gatgacctcg tcttgagatc cttaacctgc aaagaccctt ttccaaaaa aaggttatgc 46260
tcacagattc taggccttaa gacatgggtg tatctttctg gggggcacta tccaaccctt 46320
tatacaatga aagacgggaa gagggccagg tgtggttagt cacgcctgta atctcagcac 46380
tttaggaagc tgaagcggga ggatcacttg agcccaggag ttacaagta gctaggcaac 46440
atgatgagac ccattttcta caaaaagtga aaaaaaaaaa aaaaaaaaaa aagccagggtg 46500
tggtggctca cacctgtaat ccagcactt tgggaggctg aggcaggcag atcacgaggt 46560
caggagattg agaccatcct ggctaacacg gtgaaacccc gtcttacta aaaatacaaa 46620
aaattatggc cgggcgcagt ggctcccgcc tgtaatcca gcactttggg aggccgaggt 46680
gggtgaatta caaggtcaag agatcgagac catcttggct aacacggtga aaccccatca 46740
agatcacaag gtcaagagat ggagaccatc ctggctaaca cggtgaaacc ccgtctctac 46800
taaaaataca aaaaattagc cgggcatggt agcgggcgcc tgtagtccca gctgctcggg 46860
aggctgaggc aggagaatgg cgtgaacccg ggaggcggag ctgacggtga gccgagatcg 46920
ctccatgcca ctgcactcca gcctgggtga cagagtgaga ctccgtctca aaaaaaaaaa 46980
aaaaaaaaa aaaaaaagaa aattagccag gcacagtggc aggtgcctat tgtcccagct 47040
acttgggagg ctaaggcagg agaattggcat gaacccggga ggtggagttt gcagtgagcc 47100
gagatcatgc cactgcgctc cagcctgggc gatagagcaa gactctgtct caaaaaaaaaa 47160
agccaggcat ggtggtgcat gcctgtagt ccagctactc aagaggctga ggcaggaggg 47220
ttgtcgacc cacggagatc aaggctacag tgagccatga tcgcaccact gccctccagc 47280
ctgggtgaca gagtgtgacc ctgtctcaa gtaagtaa ataggaggag acaagtgggc 47340
agttcagact gatggtatgg gcacagtaga gactggtgca gacaggctgg cctgtgatgt 47400
caagcaactt ctgtaattgt ttccggcatc catttgtgtg tcaatttccg tgtcagtagg 47460
aagactctgt aggttgccaa gaggaataag tgggaggatc ctcccagaga ggccgggcct 47520

gcaggagggc cagttctcat gatttctcat ttggccccta cctccaggc tgtggttctg 47580
agggtgggaga cagagcctga cctctgtttg tcttgtttg tctttgcagc agcccaccca 47640
tgtgcccgtg acaatggtgg ctgctccac atctgtattg ccaagggtga tgggacacca 47700
cgggtgctcat gccagtcga cctcgtgctc ctgcagaacc tgctgacctg tggaggtagg 47760
tgtgacctag gtgctccttt ggggtgatgg acaggtaacct gattctctgc ctgctaggct 47820
gctgcctggc atccttttaa aatcacagtc cctgtggcat ccagtttcca aagctgattg 47880
tgtcttctt tgccctctt tctttctac tatgtgcatt cgggtgctatg aatttctc 47940
taagtactgc gtttctgca tctcacaat ttgttacct ttcattttc aggtagtgtg 48000
aatatttta cacttctct gagatgacat cttggctca tgtgttattt agaagtgtg 48060
cttagtttct aaagagtgg ggctttcca gctgtctc tgcaactgat ttctaattta 48120
attctactgt agtcgagag ctattttat atgatttctg ttattttaaa tgtgttgggt 48180
gtggtgtttt tgtgttatt gttttgtgt cttttgttt tgtttgctt cgtttgttt 48240
gttttgaga cagtgtctg ctctgtcact caggctggag tgcaatggcg cgatctcagc 48300
tcaccgaac ctctgcctcc cgggttcaag tgatctctt gcctcagcct cctgagtagc 48360
tggtgattaca ggtgcacgcc accataccca gctaatttt gtatttttag tagagacggg 48420
gtttcacat gttggtcagg ctggtctga actcctgacc tcgtgatccg cccacctcgg 48480
cctcccaaag tgctgggatt ataggcgtga gccactgtgc ctggccatta ggtgtgttt 48540
atcaccagc atcatgcagt ttatcttggg gaatgttctg tgtactctg aaaagaatgt 48600
ggattctgct gttgttgggt ggagtgttc agaaacatca attagatcca gttggtta 48660
agtgtcatc aggtgtctc tatcttctt tctgactgc ctgcttgagc tgcagtat 48720
tgacaggggt gtggagtctc caacttaat ggtggattg ttatttctc ctagtagttc 48780
tatcttttc tctcttcta ccttgatcc tcttctccc ctagggcttc ctggtgttag 48840
tggtgggaga gtggggtagt gaagaacctg gactttaggg ccaaagaggc cagggttcaa 48900

atcctggctc tgcacttcc cagttgagt accctggctg gtgcctgaat ctctgtgagc 48960
ctccacttcc tctctgtga aattgagagc acttacctgg caggctgtca tgggcatcaa 49020
gtaacagggc actccacctg gacctgaca cgtgatgcac aggaatgcca gctgctatgc 49080
catgggtgtg gcagtagtaa taaagtgacc atctgtatcc tcaccacagt gaagcctgtc 49140
cagggcttcc tctctatgc ccccatgcct ccaggtggcc ttggatcctg ttggttctgt 49200
gctctgctca gcgaccttcc tcccgtggga gttcctgggg gttcagcttc atctacaga 49260
cagcagcaca cactggctgt gcacccttt tttttttt tttttttt tgagatggag 49320
tctcgctttt ttgcgcagg ctgaagtga gtggtgtgat ctggctcac tgcaacctct 49380
acctcctggg ttcaagtgt tttctgcct caccctccca agtagctggg attacaggct 49440
cccaccacca cgcccggcta attttgtat ttccagtaga gatgggtgtt caccatgttg 49500
gccaggatgg tctgaactc ctgacctcag gtgatccgcc cacctcagcc tcccaaagt 49560
cagggattac aggcgtgagc caccacacc ggagtgccgg ttgttttag cagttgtct 49620
tgttctgga gagactggct cctgccagg agctcggga gtagggccgc ggggtgctgc 49680
ctcacacctc gatttggcc gtaagcagag gggacattt gtgactgtcc cctcctgag 49740
cttcccagca gtttttcc aagttacag ccaaagctc aggtggattt gcaaccaac 49800
ggtgtctgtg cacctccac tgatgccga actgccctgg ccaagaaac gggccgtcag 49860
aacgtgcac taactgcagc ctgggcctc catgccagag gccatgccct tccatccacc 49920
acccctggc ctgggccctg ggccctctg gctcgggaac tccaggcccc ttctcacgg 49980
ctcgagagac gtgtattac cgcacaggtg ctgtcattc tctgtggcc tcttccag 50040
ggagatcaca gaaggacagg gcctcactga ggtctcggac atggaccctt tgatagtggc 50100
aggagccagg ctgggcaaga ggcggccaca gtcacctcag cagtgccatc accaccgcca 50160
ttcagccctt cctgagccg ggcgcgcccc tggctctggc cccagtgtcc cagttacagc 50220
tcacaggagc ttgtgtgcc cagcggctgc ttctgattga gactcgaggt cggaggctt 50280

gggaggctga gaggctgctc ggttcacaa ctgctgaggg agacttgggc tccatctcag 50340
gtatgcccc a tgcgccctc aacctccagc caccggtcct ccgtgtcccc catggccagg 50400
cacggcttgc agacatctgt cgttggctcc tctcagccgt cgtgggctga ccctggcacg 50460
tcctcctgtg gctgagccca gtggggacag ctgcttcctt ttattaccct agaactctcg 50520
tctttgatca ggccccctcc cctatgccac acagtccctg tcaactgggt gagcccagta 50580
gtcatgggga aggcctgcgg gtccaaaca tcaaaggct tgcgtgcagc atgacagctt 50640
gaaaccgatg tttttacct tgatcagatt tcagcttggc gggggcttg ctcagcttc 50700
agtgaggcct gggccgattt ccagcatcc cctcctgagg ccagcctctg tttcctgtga 50760
ttttctgcac aaagtgggag ggaggagtcc taggaaatgg ggggccacct cgaagcctag 50820
gcctcctctg gcttctctgt gccagtgcc ccacgcttg tctctgtgc cccagccat 50880
gggactctgc tattccctga gtgtgccgc atgccagcc cgcactgagg acgtggagcc 50940
ccgaggggca ggaatggcctc catggtcaca ctaggaagt ggctccacc ctccgatgat 51000
cctctccctc ctcccttca ggcctccccc cggggggtgc ctcagccctc ctgcctgtgc 51060
ttgtcccgt ctctgcagg cgcctgggac gtgctgacag gtcctctgcc ggctcctgcc 51120
ttgctatgcg cacgctggtc accacagagg cctggccctt ctctgtagc agtcccacac 51180
ccgcaacagg tgtggctgt gaccacctgc ttctgcccc tctggtctg aggagggcgc 51240
agtgggcact caggcgtggc tgagcagatg tgtgtgccg ggaggaggaa ggactgctcc 51300
agtcagggt gaatttcca cccggagcat ttctgtgta ttggtgtag cgcctgctgc 51360
ttaaagctct gattcccagt tggcaccctt tccctctgc attgaaaaac atacggatgc 51420
atgtcttctt gcagtgaatg tgtattctcc cagcctctct tctgggttg ggctggaggt 51480
ggagcggcac acaggagccg cagcgatgga ggatgtgcgg gtgcagcacc ccgtacagca 51540
gggatgcaa acccgcgctg agtcctctc aacttctgct ttgaagcca gtcacgcat 51600
tgctgggtt ttgctgggcg gggctgctg tgatgtctc ctctgtcct ccccagagc 51660

cgcccacctg ctccccggac cagtttgcat gtgccacagg ggagatcgac tgtatccccg 51720
gggcctggcg ctgtgacggc ttcccgagt gcgatgacca gagcgacgag gagggctgcc 51780
ccgtgtgctc cgccgcccag ttccctgcg cgcgggggtca gtgtgtggac ctgcgcctgc 51840
gctgcgacgg cgaggcagac tgcaggacc gctcagacga ggcgactgt gacggtgagg 51900
ccctccccgt caaggctctg ccaagaccct ggccctgccc tccgggatac gagcttgggg 51960
ctgcctccgg cctcacagga gtaggggctc tgaaaacctt tgcttcagg gagattgcca 52020
agtctgtctt ttaggcccac caaggaaaac tctgcagttc caccatcct gtcccaccag 52080
gtagtgtggc ttgaaggcag actgtgaggg tctatctcac ctctctcat taggtcagga 52140
gtttcacaga aacctgaggc acattcaggg gtgggctgca gaggtccatg gtcacaccc 52200
tggaatatcc gccccaaaa gacagtgtg tctccactga ccagtctgtg ggatagtgt 52260
taagcctgag tggtttctat caacatgtag aatcaggagg tataaagaga ttgtctagg 52320
catcctgggc cctctctgac cagcaggatc ttctttaga tcttgacagt gaaacacatc 52380
tcttctgtgc cccctgtgag tttctttca ttcattcatt cattcattca ttcattcatt 52440
cattcattcg agacagagtc ttgctctgtc acccaggctg gaggccctg gtgtaatctc 52500
ggctcactgc aacctctgcc tccagggtc aatcgattct cctgcctcag cctcccagat 52560
agctgggatg acagggtgcg accaccatgc ctggctaatt ttgtatatt tagtagagac 52620
agggtttcac catgttgcc aggtgtgtc cgaactcctg acctcagggtg atccgcccgc 52680
ctcagcctcc caaagtgtg ggattacagg catgagccac cgcgcccgc ctgagttttc 52740
ctttatgaa ggacctgctt ggttggtgc ctgccatg ttgtcagcac catgggcccc 52800
ggactgtga ggagctgtg atgccctgc tctccagag ccaccggctc tgtagataa 52860
ttcacatga gtctggccac tgcctacgt cctcattcac aaagagcaga catttcgtag 52920
aagatgaggg cctgggagta acctccctgc atgttttct ataaaggcat agtggttaag 52980
tcttccagc tcattgacca ttggagaatt ttatggaggc ttagactag gggctggtaa 53040

actaagggcc caggggccaa atccagcctg ccacctactt ttgtaaataa agttttcttg 53100
gtgcacagcc atgcccattc attcatttgc acaatgtctg tggctgcttt catgccaaaa 53160
gcaagagaac tgagtgggta tgctggagac ctacggcctt caaagcccca gacctcacgt 53220
ctggcccttg acagacagag ctccccagc cctgctgcgc atcctggccc agcatgtgct 53280
gtgtgtgtga tticagcttg caggagccgt ggtaggaat tgtccctgtg ttgtccatt 53340
ttgcattgct atgaaggagc acctgaggcc ggtagatta tgaaggaaag aggtctgtct 53400
ggctcatggt tctgtaggca gcaccagtat ggcaccgcga tctgctcagc ttctagttag 53460
gtctcaggaa gctttgactc atgggtgaaag tcgaagcggg agcaggtgca tcacatggtg 53520
agagagggag caacggagag agagagagag cgcctctccc tctgcccctc acctgagag 53580
gagatgccag gctcctttaa gtaaccagct ccatgtgaa ctacagtga gagcccattt 53640
gctactgcgg agaggggcacc aggcctctgc tccatgacc caaacactgc ccaccaggcc 53700
ctacctcaa ccttggggtc atattttatt ctgttctatg ctatgctatg ctatgccatg 53760
ccatgccatg ccatgctatt cctattctat tatttagac agaattctgc tctgttcccc 53820
aggctggagt gcagtggcat gatcttggct cactgcaacc tccacctccc aggttcaagc 53880
gattctctg taccagctc ccgagtagct gggattacag gcacacacca ccacaccggg 53940
ctaatttttg tattttcaat agagatgggg ttaccatg ttggccaggc tggctcaaa 54000
ctctggcct caagtatcc acctacctg gcctcccaa gtgccatgat tacagatgtg 54060
agtcactgcg cccagtgagg gtcacattc cgttagatt tggaggggca gacgttggag 54120
ccatctgagc cccctcgtcc cgtctagct tctctccc tgtgccccgc ggtgctggtg 54180
gcaggccctt acgccgggtc tggctgcatg ctctgttcca gaagctttct tccctgcttg 54240
gttaccagaa aatcatccca tccattacaa ggacagggtc ccttatctc cattcccag 54300
ggcaggacac cgggggcagg gcaggtggg aactgagcaa gtctctggg ggcaggcgtg 54360
gctatggctc cctctgggtg ggcgtctggg gaggggtgga ggcagccgtc agcgccctgg 54420

cttgctcttc ctccctggcc agagacttg gccttgctgt gctcccgtgt gggctgcctg 54480
cacctccagt ggggttgct cctccccc cctccccc aagctctgt gagcaccact 54540
gccttcaca gccccactc tcgggaggcg aggtcctcg tggccattcc tgccttggc 54600
accaccccc ccaccaacct ggtagagcct tgggcgggg ctgttactcc tgcattggcg 54660
tagacctccc cacagtaggc acctgacaca tacctcctgg ggggcaggca ggaggtgcgt 54720
tgaggtctca gccctggcag tccctccct gcgtggcata ggcctcgcca cagggtcatc 54780
gagggtgggt ggagactgta ctagaccact ccccgctggt cctagaaagg gtcccatctg 54840
tctgctctct gtttggagtc cagacctgg ttgctgtgcc ctgcatggtg ggctgggggg 54900
cacctccag cctctctgag tgcattggc ctccttcag ccatctgcct gcccaaccag 54960
ttccggtgtg cgagcggcca gtgtgtctc atcaaacagc agtgcgactc ctcccccag 55020
tgtatcgacg gctccgacga gctcatgtgt ggtgagccag ctctggcac ggggaagggg 55080
cgccgggct ggggtcccc aggaacgtgg agtttagggg aggagacgtg cctttccagc 55140
ggggctgggg gctgtgtggg agactcaggc ggctgggagg ctccttcgg gaggcaggga 55200
agcctttccc agggcagcgg ccaggaggac agactgtgag ctgtgggctc ggcggctaca 55260
gagtctgcct cagtgggcgg ggctgatgt gtccagggtc ctgcagcacg caccaccca 55320
cgggacctg ctgagcagcg tctgtcaggc agcaagatta cccgagggt gcagtgtcc 55380
tgttccctgg cagcttactg tctggctgag gaggagtgt gttcacatat gcacacatgt 55440
catgtgcaca cacatgtaca tgacaacac ccacatgctc ctcaaatagc atgacctgt 55500
cagtcacgga tataggcct aggggatagg agccaagac agtcaggga gactttccag 55560
aggcagtggc tcctgaaagg ctgtctgatt caggcaggaa gggagctgag ttcagatagg 55620
aagtagcaat gagtattgt gtctggggac atggccactc ctctgctga gagggacctg 55680
ggctgagagc tcctcttta tggctgcagt cgggagagaa gtctgttggg gggagaaggg 55740
ggcttctca agggactccc tgtgccctt ggcacctc tgcaggta ggcttgaggc 55800

ctgaaggcag tgggtgggggc caccaagggt cgcctcctct gctgggcaag ttcccagtct 55860
gacgggcctg tgccgtgggc ccagctgtg ggggcgctgt tgatgcgcag ccaggcctcg 55920
ccgccagagc ccgcacgctt ccattccgct gacttcacg acgccctcag gatcgctggg 55980
ccggccctgt gggagagtga atgtggcttt tgccaaagti gagtctggag cctggaaact 56040
tccctatggg cagccttgat agtggagtgg cccaaggagc ccaccagcc gaccctgcc 56100
ctcccgtggc tgggtggcgcg caccaggggc tgcttgctt tgctcgttca ccaacatcac 56160
ctgggctggc cagggcgcgc tcacttctgc caccaccgag ggccctgggc gaaggagtga 56220
ataccaggct gccttggcag ggatgtgtg agggctgtg ggagtcggac agcggcgggg 56280
gtcagaggag gaggaggggtg caccgtgcag gctgaagggc cacgttacc tgaggttggc 56340
caggctcccc aggcctagcc tccagctcc ccactttct cccaccctc caccagtggc 56400
aaagccagcc cttcagggc gcacggtgc tgccccaag gagggcccat tccgttggg 56460
ttaatgttg ccacctctt ctgtttgtct ctggcagaaa tcaccaagcc gcctcagac 56520
gacagcccg ccacagcag tgccatcggg cccgtcattg gcatcatcct ctctcttc 56580
gtcatgggtg gtgtctattt tgtgtgccag cgcgtggtg gccagcgcta tgcgggggcc 56640
aacgggccct tccgcacga gtatgtcagc gggacccgc acgtgcccct caattcata 56700
gccccggcg gttccagca tggcccttc acaggtaagg agcctgagat atggaatgat 56760
ctggaggagg caggagagta gtctggcag ctttggggag tggagcaggg atgtgctacc 56820
ccaggccctc ttgcacatgt ggcagacatt gctaactgat cacagcttc agcctttccc 56880
actgagcctg tgcttggcat cagaatcct caacacagag gcctgcatgg ctgtagcaac 56940
ccacccttg gcactgtagg tgtggagaaa gctccttga cttgaccttc atattctagt 57000
aggacatgt ctgtgtgtc cacaaatcct catgtacct agaatgaat gtggggcg 57060
ctgggctctc tccagagctg aaggaatcac tctgtacat acagcagctt tgtcttgagt 57120
gcagctggga ttgtggctg agcagttaca attcctacgt ggcccaggca ccaggaacgc 57180

aggctgtgtt ttagatggc tgggcagccg caccgcagag ctgcaccatg ctggtttgta 57240
tcacatgggt gaccatggta tgtctaagaa ggtggagtcc ctgtgaggtc tgcaggtgcc 57300
cccacagctc caggccacct tgaggattgc ctctgcctgc ccagccctga gttccctctc 57360
ccctgtcctg tccactgtc accccaagcc ggccctattg ggagcctgtt ggatggcagg 57420
gtatagatgt aacctgattc tctctgggga gcgggggttat ctggcttctc aagagctcct 57480
aggagcccac agtgggtggca ccatcacagt cgcagcagcc ccagagaaac gcggccctgt 57540
ctgttcttgg cgtgctctgt gctgccccgc ctgggttccc tgccccagtc gcaggcccct 57600
tggaggaggt accatgtgtc tccgtttca cagatgagcc ccggggagct cactctagta 57660
gtggccagag aggctgcgg ctcaggagc ggggcacatt tccaacagga cacaccgcc 57720
tggctgagt ctctgggta gtgggagcag aggagagcgc cctatgtctg tggggcggt 57780
tggctgagcc tggaagccac ctgacctccc ccgtcccttc cctgccaggc atgcatgcg 57840
gaaagtccat gatgagctcc gtgagcctga tggggggccg gggcggggtg cccctctacg 57900
accggaacca cgtcacaggg gcctcgtcca gcagctcgtc cagcacgaag gccacgctgt 57960
acccgccggt gaggggcggg gccggggagg ggcggggcgg gatggggctg tgggcccctc 58020
ccaccgtcag tgctggccac cggaggcttc ccgggttctt gggggctgtg ccaccgcctc 58080
tgaggcatgc ttgctttctt ccttttcaa acccttctgc ttcttctt aatgacattg 58140
ttgattgtgg ataatctgaa aactacaaa aatataaag agccaaaatc tcaccaaata 58200
ccacctcta gagggtgtt tgggtccgt cagcatccag gcggccgtct gtgtccgca 58260
cggcccagcc catgatagc cgctgcacc aggcctgtct gccctctgtg agcctccca 58320
cagggttccc tccacaaaca cctgttctc ccaccaggg ctggctgctt cctggaaaac 58380
agctggatgg tttgtgcat gacagacaaa cacagggtga tttcgtggc taaaatactc 58440
cctggagctt ttggcagggt gaggggctgg ctccagctga gccacgcctt gaggaaatg 58500
actgtgagga gaataaactg ccgtgccct ccaggatcac tggggctggc tggggagAAC 58560

ccccgtttct gggagcacag tcccaggatg ccaaggcgag cttggtgccg agatgtgaac 58620
tcttgagtgt aaacagcggg ggctgacttg acatgcttgg tatgctttc atttgttct 58680
gcagctgtat gccctaagg tgagtccagc ccccttctgc ttctctggg gcctcgccag 58740
tgagccccac cttgctgggg ctggttcctc ctgcccttct gggtatccct cacatctggg 58800
gtcttgctt cttgttttct ttttctttt ttttgagac ggagtttcac tttgttgcc 58860
caggcttcag tgcaatggg tgatctctag gtcaccgca acctctgcct cccagggtca 58920
agcagttccc ctgcctcagc ctccctagta gctgggatta caggcatgtg ccaccacgcc 58980
cagctaattt tgtattttta gtagagatgg ggtttctcca tgttggtcag gctgatcttg 59040
aactccctac ctgaggtgat cgcgccacct tggcctccca aagtgctggg attacaggcg 59100
tgagccaccg cacctggcct ttttctttc ttttctttc tttttctga gacagggctt 59160
cgctctgtca cccaggctgg agtgcaatgg tgtcatcatg gctaactgca gcctctacct 59220
tctaggctca agcaatcctc ccattcagc ccctaagtag ctaggactgc acgcatgcat 59280
ccccatgccc agctaataat tacattttt gtagagatga agtttacta tattgcccag 59340
gctggtctcc aactcctgga ctgagcgat cctctgcct cgccctccc aggtgctggg 59400
attacaggcg tgagccaccg tgcctggcct ggggtattgt ctcttatgg cacctgactg 59460
tggtgggccc tgggaaggaa gtagcagaag agggttctc ttggttctt ggacagtaac 59520
tgagtgttct ggaggcccca gggcctggct ttgttaggg acaaaggga ctggtacca 59580
gaagccgaga gtttaaacac ccactgcct tcttccctgc tctgctgct gcaaccagc 59640
ttaaccagcc aggagtgtta ggaacccaag caggggcccc gagcacacag caggcagctc 59700
acgaattctc ttttctgtt ctccctggg agctgggagg atcttaatca ggcaataaga 59760
gatggcactg agcagccagc taattttta aatcacttta ttgttaacc atatgactca 59820
cccacttaa aaagggtaca gticagtggg ttttagtgta ttacagatg tgtgcaacc 59880
tcaccacagt taattttaga acatttctt gccctaaaa gaaactctgc atgaagccag 59940

ctgtttttaa attagcaaag ttattttgca tcctttaaat atatgttcat ggtacaaaat 60000
tcaaaagata cagaagagtc tgcagtccaa agagactccg ccccatgac gccaagcagg 60060
catccctggg aggcattggc tcctgcagtg tgtttctct atgtccccc aggggtcatc 60120
tgtacatatg caagcataca agagcgtgga cttgttttc caagccagaa gataattgta 60180
gatttatgtg cagttgtgag aaagagcaca gaccattta tcctctgcct ggttccccc 60240
agtgtgcct gccatcttgc atgacttcca ttctatcat aagcaagaca ctgataacga 60300
ttcttcacc ttattcagat tgacataagt gtttttgtt tgttctgag acaaacttc 60360
tctgtcacc agtgggagtg cagtggcaca atcacagctc actgcagcct caaactcctg 60420
ggctcaagcg attctcctgc ctacgtccc tcaagtagct cagatggcag gtgtgcacca 60480
tcatgccagg ctaatttta aatttttgt ggaggtagg cctcactaaa ttcctgggc 60540
tagtcttgaa ctctgagct aaagtatcc tcctgcctca gcctccaaa gtggtaggat 60600
tacaggcatg agccactgcg cctgggctga catatgtgtt ttgtaagcc cgaagatag 60660
catctgaaga gtcaacattg agccttgcct ttgctgcta acgatgtata aaagctgctg 60720
ttctgagcat ttggaggct ccagctgcc gtgtgcacc tcctagagc tctaccgtaa 60780
cccatctccg ggaggagggt ctattgttt cctcatttg caacaaggag gctgaagaac 60840
tgagcatgaa cacttggcct gggtcgttcg gttgtaggc agtggggcca ggccatcaa 60900
ctcacaacca ctttctactc tgcttcccc gcacctgaa gttgttctg tttgaggac 60960
acagccgtca cattcttgtt ggctgaacag cactcctgt caggcgtggc tgggccccca 61020
ctggagggca tcatgtcct ctctcctgct gcggttgaa cttggctgtt tcaaccactc 61080
ctgccaagtg gccctctgaa agggacagtc catctttct cagcagagg ccacactggc 61140
aaaacgtcc ctggcaccct ttctctccac ctgtctaata tagagtaaaa atggtatcat 61200
gttaagatct tcatttatat ttattttatc atgaatgatg taagcatcat tttgtgtt 61260
taagaacctt tgggcccagc gtgatggct gcagctgtaa tctcagcact ttaggaggct 61320

gagatgagcg gatcacttga ggccgggagt ttgagaccag cctggccaac atggagaaac 61380
cccgctctcta gtaaaaattt aaaaattagc cgggtatggt gatcccagct acttgggagt 61440
ctgaagcatg agaattgctt gaacatggga ggcggagggt gcagtgagcc gagatcgcg 61500
cattgcactc cagcctgggc gacagagcga gactctgtct caaaaaaaaa aaaaaaaag 61560
aaaagaaaag aaattatcaa tctcctcttt tatggcatat atatatatat atatatatat 61620
atatatatat atatatatatt tttttttt gttatgttca gaaaggcctt ccctgctctg 61680
atcataaaaa acaacttatt ttacactct ctctctttt ttttgagac agagttttgc 61740
tcctgttgcc caggctggag tgcagtggcg caatctcagc tcaactgaac ctccgcctcc 61800
cgggttggag tgattctcct gccttacctt cccgagtagc tgggattata ggcatgcacc 61860
accatgcctg gctaattttg tacttttagt agagacgggg gtttctccat gttggtcagg 61920
ctggtctcga actcgcgacc tcaggtgatc caccacctc ggctcccaa agtgctggga 61980
ttacagacgt gagccacat gccagccca cactctctt cttaacgtcc tcctccttc 62040
gtttacgtt cacatcttta attctctgg gatgtaatta gatttgatga gcaagggtggg 62100
catccagctt gttcttggc tgatggctta tgggtggcgt gaattagtcg gggctctatca 62160
ggaggcagaa actctatgag aatttgaaca gagaaagtic cgtctacagg cttattacca 62220
gggactggaa tagcagaaat tgaacagtga gatgtacaga gaacttaag aatgcaggaa 62280
taggccaggc atggtggctc acacctgtca tcccagcact ttgggagacc aaggcgggtg 62340
gatcacctga ggtcaggagt tcgagaccag cctggccaac atagtgaac cccatctcta 62400
ctaaaaatac aaaaaatta gctgggtgtg gtggcgcag cctgtaatcc cagctactcg 62460
ggaggctgag gcaggagaat cacttgaacc tgggaggcag aggttgcagt gagccgagat 62520
catgccactg tactccagcc tgggtggaag agcggaaactc tgtctgaaa aaaaaaaaaa 62580
aacaagaagt tcaactgaa gggaaaaatg ccgtattgtc ttccctttg ttatgtcacc 62640
agggcacagt ccatcccagg ctggcgctga tccacgggct ggagaggggc tgccccagaa . 62700

gaggacatgc caggaagggc ttggctggtg ttcaggagcc caggccaggt caggtcaaga 62760
ggtgttgagg ctggacggga gaggccagct aggggctcat gtaggatatg aggggtcggc 62820
ccatttcaac gtggaactg agctcttctg ctctctttc ttcttactg cattaagatt 62880
caataccgct tgggaagcag gtatttcct tcctataaag gatggttggg agcctgagtg 62940
ttgggagaaa gtgtagccgc tgagttaact acaactaggg ctgccgtcaa gcctatgggg 63000
aaagagagaa gaggacattt ggaaggagag agatcaagct gtggcaccct gggagaggac 63060
cacagaaaag aggccagtga ggggggtccc cgggtggcatc tgaaggtgtg gcccaaccag 63120
gaggtccaga ggctgccagc cgagtggccc aggagaggga acctcacagg ggctgagtgg 63180
gacccaagcc ctatccaccg tcctaaccac ccacatttct cgggaacaag acctcccaca 63240
gtggcctccc cggcagtga aatagccaaa ctggcaacat ggactttctt caactgcccg 63300
ggcgatgctg cctcagtgcc ccagggcagg caggaagctc ccacacccat tctggaatga 63360
ggggttggag gaaggctgag ctgagcaaag gacccatctc tgctctggtt ggtggggagg 63420
gagcccatta tacaagagac ccctcagggc tcagtgaggg gtgacagaga ctiggggagt 63480
agtggctgtc actgcagagg tgagagggtt tggagagaag gtacatgcct tttggccac 63540
attgagtagc acctggtagc cagttagtaa cgtgtattgg ataaacaaa gattaaacgg 63600
atgcaaaaaa aaatgttggc ttgtcttct ttaccctaaa cctcagttcc ctcaagtaga 63660
ttctgggaac accccctacc tggctggact gttgtgaagt ttaaataagc caggttaact 63720
tcacctctc cttaagaca cagctcagac actgcctcct ccaagaagcc ccctctggct 63780
tcctgtgtga atatgacggc cctctgggct ctagggtatc ttagaacaat gcttcctat 63840
ggctttggaa ccccgctgtc tcctggattg ggagcaaatg caggggagga gccacacctg 63900
actaatctct gggctctcca gcacataagt ggcataaggg cagggtgtg cccgcttcag 63960
gcacttactg aaggatgtac ttggcagagg gtaggcagcc ggcggatgag ccctcactc 64020
tccccagctg actgcgtggg cgggaaaggc ggggtcagga gaccagcct ccctgggctg 64080

tcacqacctc tgcacatcca gcccattga tcaaggggtc aatttttggg gtctgttgg 64140
gaggccagga gactctctcc aggcaattct tccaggtctt tgtgttaggg tgtgtgtgtg 64200
tgtgtgtgtg tgtgtgtgtg tgtgtgttt gttttatttt attttattat ttatttattt 64260
attttattat ttatttattt tgagacgcag tctcgctctg ttgccaggt tggaggggtg 64320
tggcatgac tcggctcact gcaagctccg cctcccgggt tcacgccatt ctctgcctc 64380
actcttctg agtagccgga ttacaggcgc acgcacatg cctggctaatt tttttgttt 64440
tttagtaga gacaggggtt cgccacgtg cccaggctgg tcttgaatcc ctggcctcaa 64500
gcgatccgcc cgcctcagcc tcccaaagt ctgggattac aggcgtgagc caccgtgcc 64560
gcccagccta ggggtacatg aaacttttt tttttttt ttgagacaga gtttactct 64620
gtctcaggc tggagtgcag tggcgtgac tcggcgtact gcaatctccg cctcccgtt 64680
caagcgattc tctgcctca gcctcccgag tagctgggat tgcaggcacg cgccaccaca 64740
cccagctaatt tttgtattt ttagtagaga cgggctttca ccatgtggga caggatggtc 64800
tcgatctcct gacctctga tccgcccgc tcagcctccg aaagtgtgg gattacaggc 64860
ctgagccacc gtgcccagcc atgatgttt gatacaggca tataacgtat aataatcaca 64920
tcagggtaaa tgatgtaacc atcacatcaa gcatttatcc ttigtgttac aaaaaaaaaat 64980
ctaattatac ttctactt attcttttt tttttttt ttgagacga gtctccctca 65040
gtcggccagg ctggagtga gtggcatgat ctcaattcac tgcaagctct gcctcctagc 65100
tctgcctcct gggttcatgc cattctctg tctcagcctc gcgagtagct gggactacag 65160
gpgcctgcca ccgtgcccgg ctaattttt ttttgtatt ttggtagag acaggggttc 65220
accgtgttag ccaggatggt ctgatctcc tgacctata atccgcccgt ctggcctcc 65280
caaagtgtg ggattacagg catgagccac cgccccagc ctatttattc ttaaatgtac 65340
aataaattat tgtgactcc agtcaccctg ctgtgtacc aaatacggat ctcttcatt 65400
ctatctaact gtatttctgt acctgtaac catctctcct ccacctcacc ccccaaacc 65460

actacccttc tcagcctctg gtaaccatcc ttctactctc tatctctatg agttcaattg 65520
tattaatttt tagctccccg gccgggcacg gtggctcacg cctgtaatcc cagcacttca 65580
ggaggctgag gcagggtgat cagcagggtca ggagttigag accagcctgg ccaacatggt 65640
ggaaccccat ctctactaaa aacacaaaaa ttagctgggc gtggtggtgg gcgctttag 65700
tcccagctac ttgggaggct gaggcaggag aatcgcttga aactgggagg cagaggttgc 65760
agtgcgcaa gattgcgcca ctgcactcca gtctgggtga cagagtaaga ttccatccc 65820
aaaaaaaaa agtttagctc ccacaaataa gtgagaacac gtgaagttc tctttctgtg 65880
cctcgcttgt ttacttaac ataatgacct ccagttccat ccacgttgtt gctttgttat 65940
aaatgacagg atcttggtca gggcagtggt ctcatgcctg taatcccagc actttgggag 66000
gctgaggtgg actgatcatg aggtcaagag atcgagacca tcctggctaa cacagtga 66060
ccccgtctct actaaaaata caagaaatta gccgggcgtg gtggtgggca cccatttccg 66120
ccccctctcg ggacgtgat gcacgacata ttacccatcc ccggaagact aatcctcccc 66180
cactctatat tgtaccttt cttttctct ccaecgatt ccccgagtaa cccgtcttcc 66240
ctccctctc ggattacgt cactttccg cttaatac gttgtccgt ccccttcccc 66300
attcgtagca ctctcactt tegtcttct accccacta tccctttcg tctctctat 66360
tcttactta ctctcccc ttcttctat acttattcc ctccgtctt cccactcgcg 66420
ctccacttt cacctagtg cctcaccta cgttgccatc tcgcccctt ttcagctctc 66480
ggcctctcac ccactgtcc tctctttac ctctctctc atctcgctca gacatctctc 66540
tagactatcc ctactttac ctctcagtc gtcttcttc tatcttctg tctccatgat 66600
cttcacgtcg ccactcttt tcgcccctt catatgtctc tcttcatgtt ctactatca 66660
tctcatgat cactatcgtt ctactactt atcactcccc tctttctca tcaattcctc 66720
tcgctattc tegtctctc ctacaaccg ccttcttgt gctatctaac tcaaccatgc 66780
ctctctact ctctcttat cgcccccca tcgcttatgc atctcttct attgcacacc 66840

cgccccctcca tcgcttatgc atcctcttct attgcacacc gccctccat cgcttatgca 66900

tcctcttcta ttgcacatcc tcttctattg cac 66933

<210> 12

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 12

ctgagcggaa ttcgtagac c 21

<210> 13

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 13

ttggtctcac gtattccgct cga 23

<210> 14

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 14

ctcgagaatt ctggatcctc

20

<210> 15

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 15

ttgaggatcc agaattctcg ag

22

<210> 16

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 16

tgtatgcgaa ttcgctgcgc g

21

<210> 17

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 17

ttcgcgcagc gaattcgcat aca

23

<210> 18

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 18

gtccactgaa ttctcagtga g

21

<210> 19

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 19

ttgtcactga gaattcagtg gac

23

<210> 20

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 20

gaatccgaat tcctggtcag c

21

<210> 21

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 21

ttgctgacca ggaattcgga ttc

23

<210> 22

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 22

cuacuacuac uactgagcgg aattcgtgag acc

33

<210> 23

<211> 32

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 23

cuacuacuac uactcgagaa ttctggatcc tc

32

<210> 24

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 24

cuacuacuac uatgtatgcg aattcgctgc gcg

33

<210> 25

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 25

cuacuacuac uagtcactg aattctcagt gag

33

<210> 26

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 26

cuacuacuac uagaatccga attcctgggc agc

33

<210> 27

<211> 45

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 27

aactggaaga attcgcgcc gcaggaattt tttttttt tttt

45

<210> 28

<211> 13

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 28

aattcggcac gag

13

<210> 29

<211> 9

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 29

ctcgtgccg

9

<210> 30

<211> 14

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 30

gtacgacggc cagt

14

<210> 31

<211> 16

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 31
aacagctatg accatg

16

<210> 32

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 32
ccaagttctg agaagtcc

18

<210> 33

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 33
aatacctgaa accatacctg

20

<210> 34

<211> 57

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 34

agctgctcgt agctgtctct ccctggatca cgggtacatg tactggacag actgggt 57

<210> 35

<211> 56

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 35

tgagacgccc ggattgagcg ggcagggata gcttattccc tgtgccgcat tacggc 56

<210> 36

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 36

agctgctcgt agctgtctct ccctgga 27

<210> 37

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 37

gccgtaatgc ggcacaggga ataagct

27

<210> 38

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 38

gagaggctat atccctgggc

20

<210> 39

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial sequence is a primer.

<400> 39

acagcacgtg ttaaagggg

20

<210> 40

<211> 163

<212> DNA

<213> Homo sapiens

<400> 40

```
actaaagcgc cgccgccgcg ccatggagcc cgagtgcgct cggcgccgggc ccgtccggcc 60
gccggacaac atggaggcag ctccgcccg gcccgcgtgg ccgctgctgc tgctgctgct 120
gctgctgctg gcgctgtgcg gctgcccggc ccccgccgcg gcc 163
```

<210> 41

<211> 419

<212> DNA

<213> Homo sapiens

<400> 41

```
gccccacagc ctgcgcgtc ctgctatttg ccaaccgccg ggacgtacgg ctggtggacg 60
cggcgaggagt caagctggag tccaccatcg tggtcagcgg cctggaggat gcggccgcag 120
tggaattcca gttttccaag ggagccgtgt actggacaga cgtgagcgag gaggccatca 180
agcagacctt cctgaaccag acggggggccg ccgtgcagaa cgtggtcatc tccggcctgg 240
tctctcccga cggcctcgcc tgcgactggg tgggcaagaa gctgtactgg acggactcag 300
agaccaaccg catcgagggt gccaacctca atggcacatc ccggaagggt ctcttctggc 360
aggaccttga ccagccgagg gccatcgctt tggaccccg ctcacgggtaa accctgctg 419
```

<210> 42

<211> 221

<212> DNA

<213> Homo sapiens

<400> 42

```
ccccgtcaca ggtacatgta ctggacagac tgggggtgaga cgccccggat tgagcgggca 60
gggatggatg gcagcaccgg gaagatcatt gtggactcgg acatttactg gcccaatgga 120
ctgaccatcg acctggagga gcagaagctc tactgggctg acgccaagct cagcttcac 180
caccgtgcc aacctggacgg ctcgttccgg taggtaccca c 221
```

<210> 43

<211> 221

<212> DNA

<213> Homo sapiens

<400> 43

```
tccctgactg caggcagaag gtggtggagg gcagcctgac gcaccccttc gccctgacgc 60
tctccgggga cactctgtac tggacagact ggcagaccgc ctccatccat gcctgcaaca 120
agcgcaactgg ggggaagagg aaggagatcc tgagtgcct atactaccc atggacatcc 180
aggtgctgag ccaggagcgg cagcctttt gtgagtgcg g 221
```

<210> 44

<211> 156

<212> DNA

<213> Homo sapiens

<400> 44

```
ttctcagtc cacactcgt gtgaggagga caatggcggc tggcccacc tgtgcctgct 60
gtccccaagc gaggcctttt acacatgcgc ctgcccacg ggtgtgcaga tgcaggacaa 120
cggcaggacg tgtaaggcag gtgaggcggg gggacg 156
```

<210> 45

<211> 416

<212> DNA

<213> Homo sapiens

<400> 45

```
ctccacagga gccgaggagg tgctgctgct ggcccggcgg acggacctac ggaggatctc 60
gctggacacg ccggacttca ccgacatcgt gctgcagggtg gacgacatcc ggcacgccat 120
tgccatcgac tacgaccgc tagagggcta tgtctactgg acagatgacg aggtgcgggc 180
catccgcagg gcgtacctgg acgggtctgg ggcgcagacg ctggtcaaca ccgagatcaa 240
cgaccccgat ggcacgcggg tcgactgggt ggcccgaac ctctactgga ccgacacggg 300
cacggaccgc atcgagggtga cgcgcctcaa cggcacctcc cgcaagatcc tgggtcgga 360
ggacctggac gagccccgag ccacgcact gcaccccggt atggggtgta acgggc 416
```

<210> 46

<211> 198

<212> DNA

<213> Homo sapiens

<400> 46

```
ttcttctcca gcctcatgta ctggacagac tggggagaga accctaaaat cgagtgtgcc 60
aacttgatg ggcaggagcg gcgtgtgctg gtcaatgcct ccctcgggtg gcccaacggc 120
ctggccctgg acctgcagga ggggaagctc tactggggag acgccaagac agacaagatc 180
gaggtgaggc tcctgtgg 198
```

<210> 47

<211> 244

<212> DNA

<213> Homo sapiens

<400> 47

```
ccgtcctgca ggtgatcaat gtgatggga cgaagaggcg gaccctcctg gaggacaagc 60
tccgcacat ttccgggttc acgtgtgctg gggacttcat ctactggact gactggcagc 120
gccgcagcat cgagcgggtg cacaaggcca aggccagccg ggacgtcatc attgaccagc 180
tgcccgaacct gatggggctc aaagctgtga atgtggccaa ggctcgtcggg gagtccgggg 240
ggtc 244
```

<210> 48

<211> 313

<212> DNA

<213> Homo sapiens

<400> 48

```
gttcgcttcc aggaaccaac ccgtgtgcgg acaggaacgg ggggtgcagc cacctgtgct 60
tctgcacacc ccacgcaacc cgggtgtggc gccccatcgg cctggagctg ctgagtgaca 120
tgaagacctg catcgtgcct gaggcctttt tggctttcac cagcagagcc gccatccaca 180
ggatctccct cgagaccaat aacaacgacg tggccatccc gtcacgggc gtcaaggagg 240
cctcagccct ggactttgat gtgtccaaca accacatcta ctggacagac gtcagcctga 300
```

aggtagcgtg ggc

313

<210> 49

<211> 255

<212> DNA

<213> Homo sapiens

<400> 49

```
cctgctgcca gaccatcagc cgcgccttca tgaacgggag ctcggtggag cacgtggtgg 60
agtttggcct tgactacccc gagggcatgg ccgttgactg gatgggcaag aacctctact 120
gggccgacac tgggaccaac agaatcgaag tggcgcggct ggacgggcag ttccggcaag 180
tcctcgtgtg gagggacttg gacaaccga ggctcgtggc cctggatccc accaaggggt 240
aagtgtttgc ctgtc                                     255
```

<210> 50

<211> 210

<212> DNA

<213> Homo sapiens

<400> 50

```
gtgccttcca gctacatcta ctggaccgag tggggcggca agccgaggat cgtgcggggc 60
ttcatggacg ggaccaactg catgacgctg gtggacaagg tggccggggc caacgacctc 120
accattgact acgctgacca ggcctctac tggaccgacc tggacaccaa catgatcgag 180
tcgtccaaca tgctgggtga gggccgggct                                     210
```

<210> 51

<211> 352

<212> DNA

<213> Homo sapiens

<400> 51

```
gtgttcatgc aggtcaggag cgggtcgtga ttgccgacga tctccgcac ccgttcggtc 60
tgacgcagta cagcgattat atctactgga cagactgaa tctgcacagc attgagcggg 120
```

ccgacaagac tagcggccgg aaccgcaccc tcatccaggg ccacctggac ttcgtgatgg 180
acatcctggt gtccactcc tcccgccagg atggcctcaa tgactgtatg cacaacaacg 240
ggcagtgtgg gcagctgtgc ctggccatcc ccggcggcca ccgctgcggc tgcgcctcac 300
actacaccct ggaccccagc agccgcaact gcagccgtaa gtgcctcatg gt 352

<210> 52

<211> 225

<212> DNA

<213> Homo sapiens

<400> 52

gcctctcta cgcccaccac cttcttctg ttcagccaga aatctgcat cagtcggatg 60
atccggagc accagcacag cccggatctc atcctgcccc tgcattggact gaggaacgtc 120
aaagccatcg actatgaccc actggacaag ttcattact ggggtggatgg gcgccagaac 180
atcaagcgag ccaaggacga cgggacccag gcaggtgccc tgtgg 225

<210> 53

<211> 235

<212> DNA

<213> Homo sapiens

<400> 53

ctttgtctta cagccctttg tttgacctc tctgagccaa ggccaaaacc cagacaggca 60
gccccacgac ctacgcatcg acatctacag ccggacactg ttctggacgt gcgaggccac 120
caataccatc aacgtccaca ggctgagcgg ggaagccatg ggggtggtgc tgcgtgggga 180
ccgcgacaag cccagggcca tcgtcgtcaa cgcggagcga gggtaggagg ccaac 235

<210> 54

<211> 218

<212> DNA

<213> Homo sapiens

<400> 54

ccaccctccc gcaggtacct gtacttcacc aacatgcagg accgggcagc caagatcgaa 60
cgcgcagccc tggacggcac cgagcgcgag gtcctcttca ccaccggcct catccgccct 120
gtggccctgg tggtggaaca cacactgggc aagctgttct ggggtggacgc ggacctgaag 180
cgattgaga gctgtgacct gtcaggtacg cgccccgg 218

<210> 55

<211> 234

<212> DNA

<213> Homo sapiens

<400> 55

ggctgcttgc agggggccaac cgcctgaccc tggaggacgc caacatcgtg cagcctctgg 60
gcctgacat ccttggaag catcttact ggatcgaccg ccagcagcag atgatcgagc 120
gtgtggagaa gaccaccggg gacaagcgga ctgcatcca gggccgtgc gccaccta 180
ctggcatcca tgcagtggag gaagtcagcc tggaggagt ctgtacgtgg gggc 234

<210> 56

<211> 157

<212> DNA

<213> Homo sapiens

<400> 56

ttgtctttgc agcagccac ccatgtgcc gtgacaatgg tggctgctcc cacatctgta 60
ttgcaagggt tgatgggaca ccacgtgct catgccagt ccacctgtg ctctgcaga 120
acctgctgac ctgtggaggt aggtgtgacc taggtgc 157

<210> 57

<211> 272

<212> DNA

<213> Homo sapiens

<400> 57

gttctctct gtccctcccc cagagccgcc cacctgctcc ccggaccagt ttgcatgtgc 60

cacaggggag atcgactgta tccccggggc ctggcgctgt gacggcttcc ccgagtgcga 120
tgaccagagc gacgaggagg gctgccccgt gtgctccgcc gccagttcc cctgcgcgcg 180
gggtcagtgt gtggacctgc gcctgcgctg cgacggcgag gcagactgtc aggaccgctc 240
agacgaggtg gactgtgacg gtgaggccct cc 272

<210> 58

<211> 134

<212> DNA

<213> Homo sapiens

<400> 58

tctccttgca gccatctgcc tgcccaacca gttccggtgt gcgagcggcc agtgtgtcct 60
catcaaacag cagtgcgact ctttccccga ctgtatcgac ggctccgacg agctcatgtg 120
tggtgagcca gctt 134

<210> 59

<211> 274

<212> DNA

<213> Homo sapiens

<400> 59

gtttgtctct ggcagaaatc accaagccgc cctcagacga cagcccggcc cacagcagtg 60
ccatcggggc cgtcattggc atcatctct ctctcttcgt catgggtggt gtctattttg 120
tgtgccagcg cgtgggtgtgc cagcgctatg cggggggccaa cggggcccttc ccgcacgagt 180
atgtcagcgg gaccccgac gtgcccctca attcatagc cccgggcggt tcccagcatg 240
gccccttcac aggtaaggag cctgagatat ggaa 274

<210> 60

<211> 164

<212> DNA

<213> Homo sapiens

<400> 60

cttccctgcc aggcacgcga tgcggaaagt ccatgatgag ctccgtgagc ctgatggggg 60
gccggggcgg ggtgcccctc tacgaccgga accacgtcac aggggcctcg tccagcagct 120
cgtccagcac gaaggccacg ctgtaccgcg cggtaggggg cggg 164

<210> 61

<211> 130

<212> DNA

<213> Homo sapiens

<400> 61
ttggctctcc tcagatcctg aaccgcgcgc cctccccggc cacggacccc tccctgtaca 60
acatggacat gtctactct tcaaacattc cggccactgc gagaccgtac aggtaggaca 120
tcccctgcag 130

<210> 62

<211> 496

<212> DNA

<213> Homo sapiens

<400> 62
tcaaacattc cggccactgc gagaccgtac aggccttaca tcattcgagg aatggcgccc 60
ccgacgacgc cctgcagcac cgacgtgtgt gacagcgact acagcgccag ccgctggaag 120
gccagcaagt actacctgga ttgaactcg gactcagacc cctatccacc cccaccacg 180
ccccacagcc agtacctgtc ggcgaggagc agctgcccgc cctcgcccgc caccgagagg 240
agctacttcc atctcttccc gcccctccg tccccctgca cggactcatc ctgacctcg 300
ccgggccact ctggcttctc tgtgccctg taaatagttt taaatatgaa caaagaaaaa 360
aataatattt atgatttaaa aaataaatat aattgggatt taaaaacat gagaaatgtg 420
aactgtgatg gggtagggcag ggctgggaga actttgtaca gtggagaaat attataaac 480
ttaattttgt aaaaca 496

THIS PAGE BLANK